Acta Genetica et Statistica Medica

In association with

Otto L. Mohr
Professor of Anatomy, Oslo

Tage Kemp
Professor of Human Genetics, Copenhagen

edited by:

Gunnar Dahlberg
Head of the State Institute of Human Genetics and Race Biology, Uppsala

Vol. V 1955 No. 2

INDEX

BASEL (Switzerland) S. KARGER NEW YORK

Printed in Switzerland

"*Acta Genetica et Statistica Medica*" is issued quarterly. Each issue has approximately 96 pages. The annual subscription rate is Swiss frs. 48.—.

No payment is made for contributions, but 50 reprints of the article will be sent to the author free of charge. Extracopies, if desired, will be supplied at a special rate. The cost of the engravings will be borne by the publishers, provided the figures and graphs are suitable for reproduction and do not exceed a reasonable number. Otherwise the author, after due notification, will be charged with the additional cost. Articles will be printed in either English, French or German, with summaries of about 10 lines. As a rule, only original papers can be accepted.

All manuscripts should be addressed to Professor *Gunnar Dahlberg*, State Institute of Human Genetics and Race Biology, Uppsala (Sweden). Corrected proofs, review copies as well as enquiries concerning subscriptions and notices, should be sent to the publishers, *S. Karger Ltd.*, Arnold Böcklinstrasse 25, Basle (Switzerland).

«*Acta Genetica et Statistica Medica*» paraît en fascicules trimestriels d'environ 96 pages. Le prix de l'abonnement annuel est de frs. suisses 48.—.

Les collaborateurs reçoivent à titre d'honoraires pour leurs travaux originaux 50 tirages à part gratuits. Les tirages à part supplémentaires seront facturés à un prix modéré. La maison d'édition se charge des frais de clichés à condition qu'elle reçoive des originaux se prêtant à la reproduction et dont le nombre ne dépasse pas la mesure strictement nécessaire. Autrement les frais supplémentaires seront, après avertissement, à la charge de l'auteur. Les travaux pourront être rédigés en langue anglaise, française ou allemande et doivent être suivis d'un court résumé d'environ 10 lignes. Ne seront acceptés en principe que travaux originaux inédits.

Tous les manuscrits sont à adresser au Prof. Dr. *Gunnar Dahlberg*, State Institute of Human Genetics and Race Biology, Uppsala (Suède). Les épreuves corrigées, les ouvrages à analyser, de même que toute correspondance concernant les abonnements et la publicité sont à adresser à *S. Karger S.A.* Editeurs, Arnold Böcklinstrasse 25, Bâle (Suisse).

«*Acta Genetica et Statistica Medica*» erscheint vierteljährlich in Heften von etwa 96 Seiten zum Jahresabonnementspreis von sFr. 48.—.

Mitarbeiter erhalten für ihre Originalarbeiten an Stelle eines Honorars 50 Sonderdrucke kostenfrei; weitere Separata gegen mäßige Berechnung. Die Kosten der Clichés übernimmt der Verlag, soweit reproduktionsfähige Vorlagen geliefert werden, und die Zahl der Abbildungen das notwendige Maß nicht überschreitet. Andernfalls gehen die Mehrkosten zu Lasten des Autors und werden vorher mitgeteilt. Die Arbeiten können in englischer, französischer oder deutscher Sprache eingereicht werden und sind mit einer kurzen, etwa zehnzeiligen Zusammenfassung zu versehen. Es werden grundsätzlich nur unveröffentlichte Originalarbeiten angenommen.

Alle Manuskripte sind zu richten an Prof. Dr. *Gunnar Dahlberg*, State Institute of Human Genetics and Race Biology, Uppsala (Schweden). Korrigierte Fahnen, Rezensionsexemplare sowie Zuschriften betreffend Abonnemente und Inserate sind an den Verlag *S. Karger AG.*, Arnold Böcklinstraße 25, Basel (Schweiz) zu senden.

To Heinz Karger on November 10th, 1955

Heinz Karger

On the occasion of Dr. Heinz Karger's 60th birthday, I wish him the very best of luck. As Editor of the «Acta Genetica et Statistica Medica» I wish to thank him warmly for all his help and untiring efforts in the interest of the Journal. Finally I personally wish to thank him for his friendship which, I sincerely hope, will continue for many years to come.

GUNNAR DAHLBERG

UPPSALA

Nordling, C. O.: Acta genet. *5*, 93–104, 1955

From Svenska Kommittén för Geografisk Patologi, Karolinska Sjukhuset, Stockholm 60, Sweden

EVIDENCE REGARDING THE MULTIPLE MUTATION THEORY OF THE CANCER-INDUCING MECHANISM

By C. O. NORDLING, Stockholm

Introduction

According to the new theory on the cancer-inducing mechanism (*Nordling* [1953]), cancer can be induced through a series of irreversible genetic mutations occurring in one and the same somatic cell. The author of this article asserts that malignant tumours in man develop more or less as follows:

In the course of events some of the cells of the body are struck by mutations, chiefly in tissues exposed to exogenic mutagens. If the cells propagate, as they usually do in young, and sometimes in adult tissues (especially in the epidermis), there will in some cases be small colonies of cells carrying the same mutation—descendants from one mutant. Some of these mutated cells will mutate again in other loci, and certainly the twice mutated cells can also multiply. Thus this process can continue in several steps until cells are produced which contain a number of mutations and differ markedly from the original, normal cell race. Even these multiple mutants may propagate and form whole colonies. By pure chance some of these colonies will be composed of cells carrying mutated genes that make the cells react with growth (net proliferation) upon a stimulus produced by themselves. Normal tissues, and even some benign tumours, seem to stop growing after having reached a volume that causes a decreasing supply of growth stimulus. These supposed colonies of multi-mutated cells will, however, grow incessantly and independently, since their supply of growth stimulus will enlarge as the colony increases in

volume. The cells of such a colony will show all kinds of properties similar to those of other cells exposed to growth stimulus, e.g. cells from a healing wound or an embryo. In other words, such colonies will be identical with malignant tumours.

It should be pointed out that this theory does not claim to explain the development of all kinds of neoplasms that can be produced under special experimental conditions; not even all kinds of tumours in man. The theory has been put forward in order to interpret the bulk of human malignant tumours—except tumours of the nervous system—developing under "normal" conditions. In experiments and special circumstances it is probably possible to produce tumours also in various other ways, as well as, incidentally, in the "normal" way.

Facts supporting the above theory

I am aware of the following facts regarding the present theory: —

1. A cancerous tumour generally develops from such a small focus that there are reasons to assume that the whole tumour is descended from a single cell. Indeed, *Butenandt* [1949] asserts that cancer originates from some hundreds of independently derived cancer cells, and *Fischer* and *Holloman* [1951] launched a hypothesis that seven such cells are required. It is known, e.g. from the works of *Yoshida* [1944, 1952] and *Ishibashi* [1950], that a single cancerous cell is sufficient for the development of a transplanted tumour. Therefore there is no need and hardly any reason to assume a plurality of independently derived malignant cells as the initiation of cancer.

2. The properties and qualities of cancerous cells are all inherited from one cell-generation to the next, irrespective of whether they are specific to the tumour cells or common to the cells of the tissue in which the tumour originates. It is commonly held (*Engelbreth-Holm* [1949]) that the cancerous cells cannot be reverted to normal ones. That is to say, that the progeny of cancerous cells are never normal cells.

3. All the cancer cells are descendants from cells of the affected tissue, and thus from originally normal cells. (At least the egg cell must be regarded as a normal one as compared to cancer cells.)

Thus the first cancer cell must be either infected (with some external substance), or mutated, or both. The evidence of tumours produced by the single stimulus of radiation seems to exclude the possibility of sole infection, and the hereditary character of the

malignant property in the cells is incompatible with the concept of infection.

4. Cancer seems to be able to originate in at least three different ways: caused by radiation, caused by certain groups of chemicals, and without any known external cause, so-called spontaneous cancer. It is conspicuous that mutations can be produced in the same three general ways, and it can hardly be dismissed as a random coincidence that even the kinds of radiation and some of the chemicals (*Demerec* [1948]) are the same in both cases. This relation was first pointed out by *Muller* [1927] and has been much dealt with, especially by *Bauer* [1928, 1949]. Much evidence on this point has been accumulated, as becomes clear from surveys e.g. by *Danneel* [1953], *Darlington* [1948], *Fardon* [1953] and *Strong* [1949]. Since, as *Muller* [1951] states, mutations of the genes are produced by radiation with the same frequency in the somatic cells as in the germ cells, it would be strange indeed if there were not among the diversity of mutations also such as could cause independent growth. Incidentally, the way in which cancer occurs in experiments is characterised, according to *Iversen* and *Arley* [1950], by "stepwise, quantum like hits" or exactly the same as what characterises gene mutations. Already some 40 years ago it appeared logical to *Tyzzer* [1916] to regard a tumour as a manifestation of somatic mutation—not to mention *Boveri*'s even older theory.

5. It has not, however, been demonstrated whether these cancerous mutations take place in the genes of the chromosomes or in so-called plasmagenes of the cells—or in both. There seems to be some evidence that points in both directions. According to *Darlington* [1948], it cannot be the question of nuclear mutations, since there is a difference between the relative efficiency of the various agents as mutagens, and their efficiency as carcinogens. Such differences, however, are also to be found between the relative efficiency of the same carcinogen in different species, as stated for instance by *Engelbreth-Holm* [1949]. Even the various tissues of the same animal may react differently to one single carcinogen. Also the mutation-producing efficiency has been proved to vary between different species, nay between different genes in cells of the same species (*Demerec* and *Hanson* [1951]). It must also be remembered that, e.g. the crown-galls of plants have been clearly demonstrated to be caused by certain properties of plasmagenes, as was shown by *Brown* [1952]. Since the malignant capacities of cancer cells have never been diluted in the same way as the capacity of abnormal growth in crown-galls,

it seems to be more probable that the carcinogenity of animal cells is not inherent in the plasma, but in the genes of the nucleus.

6. The finding that mutations, and probably chiefly gene mutations play a part in the producing of cancer is in no way sufficient to explain carcinogenesis in man. It is quite clear that the development of cancer advances in two different steps. Some agents, called cocarcinogens, can—despite their inability to *produce* cancer—*contribute* effectively to the development of tumours when applied after a real carcinogen. This has been demonstrated many times, e.g. by *Berenblum* and *Shubik* [1947].

7. Experiments with croton oil as cocarcinogen do not imply any explanation of the way in which these agents operate. It is, however, widely supposed that a relatively high number of cancerous cells are required in order to produce an independent tumour. Thus it might be expected that any agent causing proliferation will act as a cocarcinogen. As shown e.g. by *Linell* [1947], mechanical trauma is an efficient cocarcinogen, which seems to demonstrate that the second necessary step in cancer development is nothing but the proliferation of the first tumour cell to a number sufficient for independent growth. This number might well be in the order of 100 to 1000, as suggested by *Butenandt* [1949]. It seems to be useful to refer to the first step as "initiation of malignant cell change", and to the second as "promotion of malignant cell proliferation", as has been done in recent works.

8. This two-step theory can, to a certain extent, explain the period of latency that is often found between the application of a carcinogen and the outbreak of tumours. It can also explain why malignant tumours are most likely to be found in tissues which have undergone cell proliferation in connection with puberty, healed wounds, hyperplasias, benign tumours, etc. But it can hardly explain the tendency for one form of cancer to be succeeded by others more malignant in character, as noticed by *Rous* and *Kidd* [1941] and *Snell* [1949], nor the existence of so-called precancerous states, for instance in skin exposed to X-rays. As a matter of fact, there are also all kinds of intermediary states between hyperplasia and benign tumours, as well as between benign tumours and malignant ones. Also other observations concerning the slowness of cancer development, e.g. *Caspersson*'s and *Santesson*'s results [1942] point against the hypothesis of one somatic mutation as the first step. This implies that the first step in cancer development—the mutation mechanism—can be in turn divided into a number of sub-steps, each being one single mutation.

9. This, says *Muller* [1951], "would cause the frequency of such growths (cancer) to depend upon some exponent of the dose higher than one, rather than upon the dose itself in the simple proportionate way characteristic of individual gene mutations. In that case, too, the time element would constitute an influential factor unlike what is found to be the case in ordinary mutation production." If the carcinogen dose is spread out throughout life, the effect of the time element would mean that the frequency of cancer must depend upon an exponent of the exposition time, i.e. age.

10. As a matter of fact, the cancer mortality frequency curve in males between the ages of 30 and 80 years depends approximately upon the sixth exponent of age, implying (if the time-lag between initiation and death is assumed to be short) that seven mutations are necessary for the production of the first cancerous cell (*Nordling* [1952, 1953]). This frequency curve has previously been explained in different ways. One suggestion is that the somatic resistance against this disease decreases with advancing age. However, animal experiments and metastases in man demonstrate clearly that age in itself has an insignificant effect upon the susceptibility to cancer. Also *Druckrey* [1943] has shown that time or age in itself is unimportant, the cancer incidence depending entirely upon the accumulation of a sufficient total dose of carcinogen.

Another explanation is offered by the concept of latency. Most infectious diseases have a typical period of incubation or latency, and it has been suggested that the concentration of cancer cases in old age may be due to a delay of similar character. Since the cancer producing agents probably influence the body throughout life, the frequency of cancer should be constant after an age corresponding to the duration of the period of latency. The normal spread in the duration period would result in a steep rise of the cancer frequency immediately before and after the age representing the average duration of latency. The mean duration of latency calculated in this way is found to be approximately 80 years, with a standard deviation of between one and two decades. This is certainly not at all in consistency with the facts known about cancer latency in experiments and metastases which concern a period of a few years or a decade at the most. The duration of the period of latency added to the period of time between the birth and the beginning of exposure can theoretically be calculated from the detailed form of the age frequency curve of cancer of each organ. Although this is in practice extremely

difficult for most forms of cancer owing to inadequate statistics and changing frequencies, it has been done for cancer of the stomach in England and Wales by *Stocks* [1953]. According to his calculation, the frequency of this form of cancer can be better described as depending upon the forth exponent of the part of age exceeding 18 years than as depending upon the sixth exponent of the total age. This implies that only five mutations are required for the initiation of cancer and that the sum of the periods of time from birth till the beginning of the cancer-inducing process and from the end of the same process till death is about 18 years.

It may seem plausible to suggest that the low frequency of cancer in young and middle age should be due to the natural selection operating before and during reproductive age. This suggestion, however, offers no explanation of how the supposed surviving race is constructed so as to avoid cancer in certain ages, but not in general. The problem of describing the cancer-inducing mechanism in actual mankind still remains, whether or not this mechanism is the result of natural selection.

One could also think that the age frequency curve of cancer could be explained in the same way as the curves of other diseases showing the same general form as that of cancer. Unfortunately, reliable morbidity curves are not available for most diseases, and mortality curves cannot, for obvious reasons, be substituted for morbidity curves in non-fatal diseases, as can, to a certain extent, be done in the case of cancer. Certainly, mortality curves can be compared as such, but then we must remember that the situation is more complicated as soon as mortality is numerically only a small fraction of morbidity. We will find that most diseases have mortality curves of other forms than cancer. According to *Simms* [1946], the frequency usually depends upon a variable exponent of a constant, the exponent being proportional to age, (so-called logarithmic increase). In the case of cancer the frequency depends upon a constant exponent of age as variable, and this type of curve is repeated only in the case of arteriosclerosis and causes of death depending upon arteriosclerotic degeneration, e.g. diabetes mellitus, as pointed out by *Norman* [1952]. The complete mechanism involved in the induction of arteriosclerosis has not yet been discovered, and cannot therefore be used as an alternative description of the cancer mechanism. Instead, since we have found reasons to suppose that multiple mutations are responsible for the form of the cancer frequency curve, we might infer that

multiple mutations—or other multiple random events—play a part in the induction of arteriosclerosis. This is certainly not at all impossible, since cellular alterations are involved also in the latter disease.

This two-step theory with multiple mutations as the first step can elucidate, among other things, the large differences in cancer frequency of certain organs among peoples living under apparently almost identical conditions. As the incidence does not depend upon the dose itself, but upon an exponent of the dose—probably the forth exponent—, a minute difference in dose is required in order to produce a remarkable difference in tumour frequency. Two communities using the same mutagenic stimulant would, for instance, have different results according to the manner of use. If in one of the communities the stimulant affects a certain organ for a time only by 19% longer than in the other, it would result in the former having *twice* (1.19 to the forth being 2.00) as many cases of cancer in that organ, as the latter (if only mutations caused by the said stimulant are counted). In this connection the reader may consider, e.g. that persons now 75 years old who began to smoke at 12, have had their lungs exposed to the stimulant for a period of time by 19% longer than other persons of the same age who began to smoke at 22.

Discussion

In spite of the usefulness of the multiple mutation theory, it can be criticised in many ways. Let us now consider some of the most important criticisms:

Mutations are known to occur but rarely, and the coincidence of seven various mutations must be extremely rare, and seemingly more so than the actual occurrence of tumours.—Certainly it is true that mutations are rare, but the number of cells of an individual is enormous, and one mutation in, say, thousand milliards of cells means some hundreds of mutations in the human body. We must also consider that many cancers are obviously not induced by the rare spontaneous mutations, but by mutations provoked by exogenic agents. Further, the occurrence of mutants among the somatic cells does not depend solely on the number of mutations, but on the proliferation rate of the tissue in question as well. It will, therefore, be very difficult to calculate an expected tumour frequency from the known mutation rates. It should, however, be pointed out that we need not assume that five (or seven) specific mutations are required

for the induction of a malignant tumour. Most of the mutations in question may be of different kinds in various loci, only sharing the common property of having a general deteriorating effect upon the affected cell. Perhaps some 10% of all mutated cells may be sufficiently deteriorated so as to produce—through high mortality, as suggested by *Fischer* [1930], or in other ways—some mitotic stimulus. In addition to these "negative" mutations, only one or two "positive" ones may be required,—i.e., mutations which increase the capacity of mitotic reaction upon stimuli from dead or deteriorated cells. The spontaneous frequency of the "positive" gene mutations may be in the order of one in 10 000 cells per generation (of 30 years), and the rate of "negative" mutations one in 20 cells per generation, according e.g. to figures calculated by *Muller* [1950]. With these rates it can be calculated that even spontaneous mutations, without further multiplying, would give an expected tumour frequency of at least the order of actual frequency.

The actual cancer curve does not fit in with the theoretical curve at ages below 30 years, the actual frequency being higher than the one theoretically expected.—First, it should be mentioned that the actual curve does not conform any better to other theoretical curves, such as the normal curve describing the case of a latency period of 80 years. Secondly, the relatively high incidence in youth is what would be expected considering the existence of normal proliferation in this age, with its multiplying effect upon mutated cells.

The actual curve does not fit in with the theoretical at ages above 80 either, the actual rate being lower.—Even here the normal curve can do no better. Further, people above 80 years old are selected through previous high (cancer and other) death rates in the 70's, and do not represent a random sample from the population as a whole. Probably they have been less exposed to carcinogens than the average. Finally, the cancer mortality does not at all equal the morbidity frequency in these ages, because the mortality from other diseases is high, and all those affected by cancer do not live long enough to die of it. Therefore we have every reason to expect a lower incidence in mortality statistics than the strictly mathematically calculated ones. The decreasing cancer rate with advancing age above 80, which appears in some countries, is not to be found in those with the most accurate vital statistics, e.g. the United States (white population), England and Wales and Denmark.

Some tumours, such as cancer of the lung, do not follow the

sixth-power-of-age curve even between 30 and 80.—This is certainly due to the fact that the incidence of lung cancer, and to a smaller extent also other forms, has not been static but has changed. In the case of lung cancer, the younger generations have probably been more exposed to certain carcinogens than the older ones, and the members of each cohort have apparently also been exposed increasingly with advancing age. Therefore the whole curve becomes dislocated.—It might be mentioned that the increase in exposure to carcinogen required in order to explain e.g. the five-fold increase in lung cancer frequency in 30 years among male persons of a given age in Copenhagen, as found by *Clemmesen* et al. [1953], would be about 30 per cent according to the sixth-power-of-dose hypothesis and some 50 per cent if the fourth power (five mutations) is assumed instead.

The incidence of cancer in the female genital organs, and in the prostata in the male, does not conform with the mathematical curve.—This must be due to the more or less isolated periods of high proliferation—normal or pathological—in the organs in question.

The female cancer curve, with the exception of the above mentioned organs, shows a "break" in the 40's.—Since we know that artificial hormonal changes often have a staunching effect upon the continuation of cancer, it would be expected already from this that the hormonal changes of the menopause must lower the cancer frequency. It is certainly possible that not all cancerous cell colonies reach the critical number required for irreversible continuation. Some may die out while the cells are but few, and the frequency of dying out may be increased by hormonal changes.

Malignant gliomas occur despite the absence of mitotic activity and exposition to mutagenic noxae in the nervous tissue and the glias.—Malignant tumours of the brain, and other parts of the nervous system, as registered causes of death in the Vital Statistics of the U.S. [1949], show—when related with due figures from the U.S. Census of Population [1950]—approximately the same form of frequency curve as deaths from benign tumours of the same organs. It can therefore be questioned whether or not a real difference exists between malignant and benign gliomas. As a matter of fact, there seems to be much confusion about how to distinguish in practice between malignant and benign properties in an organ (the brain) where a tumour will always prove fatal if not removed. In any case, gliomas have a statistical frequency more like that of benign tumours, and it seems advisable that their causes should be searched for in the

same direction as the causes of benign tumours in general, which have not been dealt with in this paper.

Leukemia and sarcoma do not follow the sixth-power-of-age frequency curve.—Since the mitotic activity is very high in the foetal stage, the few spontaneous mutations that occur during this period must be multiplied by great numbers. Thus there will be a "reserve" of cells that have mutated several times but still lack one or a few mutations before they can give rise to a neoplasm. In the course of life some of these cells will be struck by the remaining mutations, resulting in carcinomas, sarcomas and leukemias. The carcinoma rate in man would probably be no higher than the rates of sarcoma and leukemia, if only spontaneous mutations, but no induced ones, were operating throughout life.

Cancer is rare in the duodenum and small intestine in spite of the extremely high proliferation rate of these organs.—Mitotic superactivity must not necessarily be connected with a high cancer incidence, since e.g. an epithelium which regenerates swiftly may thrust away the outer, exposed cells before they have been able to accumulate any considerable number of mutations.

Hereditary factors play a part in the production of tumours.—The susceptibility to mutations varies between individuals, and the differences are certainly inherited. Abnormal proliferation in certain organs may be an inherited property, and will always increase the risk of cancer in that specific organ.

The risk of secondary cancer (not metastases) among persons once hit by the disease is higher than the risk of cancer in general, as noticed by *Norman* [1952].—This may be due to the hereditary factors mentioned, or to the way of living. For instance, those who use some stimulant, in most cases expose more than one organ to it.

The multiple mutation theory is contradictory to the virus theory in favour of which there is much evidence showing that tumours can be transmitted through cell-free extracts.—This may be a contradiction in terms but not in fact. It is known that such cancer-virus can be produced through the mutation of normal cell constituents, and it is also known that both virus particles and genes are intra-cellular auto-catalyst molecules. Consequently there is a possibility that the difference between gene and virus can be reduced through mutations, and it is not impossible that mutated genes could, under certain circumstances, behave as virus particles are

known to do. To avoid the apparent contradiction, the theory could be expressed in more general terms, e.g. as follows:—

"Cancer is the result of a number of certain sudden atomic rearrangements in certain auto-catalyst molecules (or duplicants) normally present in animal cells."

Conclusion

The multiple mutation theory offers a means of calculating the difference in dose of carcinogen from the difference in cancer frequency or vice versa, and it may thus elucidate the reason why minute divergences in the manner of living frequently result in considerable disparity in cancer incidence. The validity of the theory should be tested not only by further experiments but also by gathering statistical material with bearing upon the dose-frequency relationship.

Summary

The multiple mutation theory on cancer involves three different parts, viz. 1) that primary cancer evolves from one initial cancerous cell having been multiplied in some way or other above a critical number, 2) that initial cancer cells are produced from normal ones by means of genetic mutations, and 3) that not one but a number (five to seven) of different mutations in the same cell are required for the production of a cancerous cell.

This theory is supported by the following facts: 1) the correlation between mutagens and carcinogens, 2) the correlation between cell proliferation and cancer incidence, 3) the kind of mathematical relation between cancer frequency and age in man, and 4) the stepwise increase of the malignancy of tumours.

Résumé

La théorie selon laquelle le cancer provient de mutations multiples comporte trois parties différentes, c'est-à-dire 1. que le cancer primaire se développe d'une cellule cancéreuse initiale qui d'une façon ou d'autre s'est multipliée jusqu'à un nombre critique, 2. que des cellules cancéreuses initiales se produisent de cellules normales par mutations génétiques, et 3. que pas une seule mais un certain nombre (cinq à sept) de mutations de la même cellule sont nécessaires pour qu'une cellule cancéreuse se forme.

Les principaux faits à l'appui de cette théorie sont: 1. La corrélation entre mutagens et carcinogens, 2. la corrélation entre la proli-

fération de cellules et l'apparition du cancer, 3. la nature du rapport mathématique entre la fréquence du cancer et l'âge de l'homme, et 4. l'augmentation pas à pas de la malignité des tumeurs.

Zusammenfassung

Die Theorie der multiplen Mutation für den Krebs umfaßt drei verschiedene Teile, nämlich 1. daß sich primärer Krebs aus einer anfänglichen, krebsartigen Zelle entwickelt, welche auf die eine oder die andere Weise über eine kritische Anzahl hinaus vervielfältigt worden ist, 2. daß anfängliche Krebszellen aus normalen Zellen auf dem Wege genetischer Mutationen entstehen und 3. daß nicht eine, sondern eine Anzahl (fünf bis sieben) von verschiedenen Mutationen in derselben Zelle zur Erzeugung einer Krebszelle erforderlich sind.

Zur Erhärtung der Theorie lassen sich, neben anderen, folgende Faktoren anführen: 1. die Korrelation zwischen Mutagenen und Carcinogenen, 2. die Korrelation zwischen Zellproliferation und dem Auftreten von Krebs, 3. die Art der mathematischen Relation zwischen der Frequenz von Krebs und dem Alter beim Menschen und 4. die schrittweise Steigerung der Bösartigkeit der Geschwülste.

REFERENCES

Bauer, K.H.: Mutationstheorie der Geschwulstentstehung, Berlin 1928; Das Krebsproblem, Springer, Berlin/Göttingen/Heidelberg 1949. – *Berenblum, I.* and *P. Shubik:* Brit. J. Cancer *1*, 379, 1947. – *Braun, A.C.:* Sci. Amer. *186*, 66, 1952. – *Butenandt, A.:* Verh. dtsch. Ges. inn. Med. 55. Kongr. in Wiesbaden. 1949. – *Caspersson, T.* and *L. Santesson:* Acta radiol. scand. Suppl. 46, 1942. – *Clemmesen, J., A. Nielsen* and *E. Jensen:* Acta Un. int. Cancr. *9*, 603, 1953. – *Danneel, R.:* Z. Krebsforsch. *59*, 167, 1953. – *Darlington, C. D.:* Brit. J. Cancer *2*, 118, 1948. – *Demerec, M.:* Brit. J. Cancer *2*, 114, 1948. – *Demerec* and *Hanson:* Cold Spr. Harb. Symp. quant. Biol. *16*, 215, 1951. – *Druckrey, H.:* Klin. Wschr. *22*, 532, 1943. – *Engelbreth-Holm, J.:* Nyere synspunkter i kraeftforskningen. Medicinska Framsteg. 1. Gleerups Förlag, Lund 1949. – *Fardon, J.C.:* Science *117*, 441, 1953. – *Fischer, A.:* Gewebezüchtung (Müller und Steiniger) München 1930. – *Fisher, J. C.* and *J. H. Holloman:* Cancer *4*, 916, 1951. – *Ishibashi, K.:* Gann. *41*, 1, 1950. – *Iversen, S.* and *N. Arley:* Acta path. microbiol. scand. *27*, 773, 1950. – *Linell, F.:* Acta path. microbiol. scand. Suppl. 71, 1947. – *Muller, H. J.:* Science *66*, 84, 1927; Amer. J. hum. Genet. *2*, 111, 1950; Sci. in Progr. *7*, 130, 1952. – *Nordling, C.O.:* Nord. Med. *47*, 817, 1952; Brit. J. Cancer *7*, 68, 1953. – *Norman, A.:* Cancer *5*, 581, 1952. – *Rous, P.* and *J.G. Kidd:* J. exp. Med. *73*, 365, 1941. – *Simms, H.S.:* J. Gerontol. *1*, 13, 1946. – *Snell, G. B.:* Proc. nat. Cancer Conference. 1949. – *Stocks, P.:* Brit. J. Cancer *7*, 407, 1953. – *Strong, L.C.:* Yale J. Biol. Med. *21*, 293, 1949. – *Tyzzer, E.E.:* Cancer Res. *1*, 125, 1916. – United States Census of Population 1950. – Vital Statistics of the United States 1949. – *Yoshida, T.:* Proc. imp. Acad. Tokyo *20*, 611, 1944; J. nat. Cancer Inst. *12*, 947, 1952.

Herdan, G.: Acta genet. *5*, 105–115, 1955

Department of Preventive Medicine, University of Bristol

A NEW STATISTICAL METHOD FOR COMPARING PAPER ELECTROPHORESIS DATA FROM NORMAL AND PATHOLOGICAL SERA

By G. HERDAN

It is due to *Tiselius* [1930, 1937] that we know how to separate serum proteins having different velocities in an electric field. The adaptation of the method to clinical routine work in the form of Paper Electrophoresis is due to *Grassman* and *Hannig* [1952, 1951].

The principle of the paper electrophoresis procedure can be described as follows. Under the influence of an electric field, serum constituents of different molecular weight and different z-potential will move with different velocities and, therefore, accumulate in different sections of the filter paper strip.

The photometric evaluation of the distribution of proteins on the filter paper yields the light-extinction curve from which the proportions of the different constituents can be calculated.

The numerical results are usually given as a series of percentages of the five major constituents: Albumin, Globulin α_1, Globulin α_2, Globulin β, Globulin γ. The following table and Diagram shows the distribution for a normal serum, a nephrosis serum and a liver cirrhosis serum.

Table 1

	Normal	Nephrosis	Cirrhosis
Albumin	60.1 %	33.7 %	29.3 %
Globulin α_1	4.4 %	5.7 %	4.1 %
Globulin α_2	11.5 %	27.2 %	6.9 %
Globulin β	13.3 %	22.4 %	14.7 %
Globulin γ	10.3 %	12.0 %	45.0 %

In clinical work, the comparison between the pathological and the normal serum is, as a rule, carried out by mere inspection of the two curves or of the two series of percentages. This, in a way, seems somewhat inadequate considering the great amount of work that is needed for obtaining such graphical and numerical results. Considering all the trouble one takes during the determination to avoid experimental and other errors, it would seem rather an anti-climax to depend in the end only upon a subjective judgment as to whether the two graphs or the two percentage series are different or not. In this connection one must not forget that the determination of the proteins in the normal serum will fluctuate not only because of experimental errors, but also because of biological differences between individuals, and the same applies, of course, to the pathological serum.

It would therefore seem desirable to have an objective method of deciding whether a difference between two series as observed can be attributed to a real factor—the disease in this case—or whether it can still be accounted for by experimental errors and by biological variations between individuals.

The Chi-square method is not suitable for this purpose because the data for that method to be used must be in the form of numbers, that is absolute frequencies, and not in that of relative frequencies or percentages. It is feasible to express the electrophoresis results in absolute numbers, for instance, of the weight units (milligrammes) of the different constituents. This would formally make it possible to apply the Chi-square test. But although the data can be made formally suitable for the Chi-square test, they are intrinsically different from the data for which this test would be legitimate. The frequencies to be compared by the Chi-square test are assumed to be subject to random sampling fluctuations only. The percentages or frequencies obtained by electrophoresis, on the other hand, are subject, apart from random sampling fluctuations, to experimental errors and the errors of the evaluation method. On the whole, they are subject to the three following types of error:

1. Fluctuations in the electrophoresis procedure itself, that is, experimental errors.
2. Errors in observation during the photometric measurements.
3. Errors occurring during the drawing-up of curves and their planimetric evaluation.

A method which is more to be recommended would be to

compare series of *average* percentages for the normal and the pathological serum. This presupposes that a number, say between 5 and 10, of determinations are available for each type of serum, preferably from different individuals. We would then proceed in the same way as we proceed with blood or bone marrow counts.

That is, for each constituent we compare the average percentages in the normal and in the pathological sera. The difference between these two average percentages is tested against the standard deviation of such differences, according to the t-test. However, this method, although theoretically correct, and practically not too difficult, has its shortcomings. One is that it presupposes already a number of determinations of normal and pathological sera and thus is not applicable to the single case. There is, however, a more important shortcoming which would attach to any of the usual significance tests when applied in this case, and that is that they refer only to the separate components of the mixture and not to the mixture itself as it is found in the body. It is now evident that any information about the *mixture as such* is more relevant to the problem of the influence of that mixture upon the disease than the information about its separate constituents. The possibility of characterising the serum proteins by a quantity which gives information about the mixture as found and as active in the body, is provided by a recent statistical technique called the statistical theory of information. That quantity is the average logarithm of the reciprocal probabilities of the various proteins in the mixture, and is called the *Entropy* because of its formal similarity with the Entropy of mixtures of gas molecules. In symbols (*Shannon* [1948])

$$H = -\Sigma\, p_i \log_2 p_i \tag{1}$$

where p_i represents the percentage of a constituent and $\log_2$ the logarithm to the base 2 and the sign Σ the symbol for summation. Since $\log_2 p = 3.322 \log_{10} p$, the formula becomes when using logarithm to the base 10

$$H = -\Sigma\, p_i\ 3.322 \log_{10} p_i = -3.322\, \Sigma\, p_i \log_{10} p_i \tag{1a}$$

For the purpose of explaining the meaning of this quantity, the basic ideas of the theory of information must be discussed however briefly.

Let us suppose we were engaged in the well-known game of guessing an object by asking questions which only may be answered by "yes" or "no". To fix our ideas, we will assume that a person

is required to guess a letter in an alphabet of 32 letters arranged in a certain order. If that person has a mathematical mind, he or she would ask if the letter was in the upper half of the 32 letters, then whether it was in the upper half of the 16 letters in which the first questions had located it and so forth. By five such questions we can fix the letter. Writing now 1 for "yes" and 0 for "no", the letter could be expressed by a symbol such as 01010, where the first digit is the answer to the first question, the second that to the second, and so on.

This suggests that the number of digits, that is, the number of "yeses" and "noes" by which the letter is fixed be taken as a measure of information. It measures not really the "*amount*" of information as it is sometimes suggested, but the *action* needed for guessing the letter. Thus we can say that for an alphabet of 32 letters, the amount of uncertainty attached to a letter is never more than 5, since with 5 guesses at the most we can ascertain which letter somebody had in mind. What we are doing in writing the guessing results in a form like 01010, or whatever the sequence of 0 and 1 may be for a given case, is to use a *dyadic* instead of a *decimal* system of notation. Just as in the decimal system any number is represented by powers of 10, so in the dyadic or binary system any number is represented as a sum of powers of 2. For instance:

$$10 = 1.2^3 + 0.2^2 + 1.2^1 + 0.2^0$$

and 10 is, therefore, written in positional notation as 1010; we then say that 10 in this notation has 4 binary digits, or 4 is the logarithm of 10 to the base 2.

We thus recognise the "information" to be the dyadic logarithm, that is the logarithm to the base 2 of the ordinal number of the letter in question.

In general, for a series of length r (maximum rank), the number of binary digits in the positional notation of r to the base 2 is an upper bound for $\log_2 r$, namely $\log_2 r$ rounded off to the next whole number (*Meyer-Eppler* [1952], *Walther* [1954]):

$$H' = (1d\ r) \tag{2}$$

H' is a measure of the action required for guessing correctly one of r different items, each having the same probability of occurrence.

Thus for our series of 32 letters, we have

$$H' = (1d\ 32) = 5$$

in words: the maximum number of "yes"—"no" decisions required for fixing a letter is 5.

But this applies only under the assumption of equal probabilities for all letters. When decoding a message we always have to take into account the fact that letters have rather unequal probabilities of occurrence. Formula (1) must therefore be transformed, if it is to give an answer as to the "information" or guessing action required in such a case.

We may write (2) as

$$H' = - \operatorname{ld} \frac{1}{r} = - \Sigma \frac{1}{r} \operatorname{ld} \frac{1}{r},$$

and since $\frac{1}{r} = \bar{p}$, the probability per symbol,

$$H' = - \Sigma \bar{p} \operatorname{ld} \bar{p} \tag{3}$$

that is, we have written it as the arithmetic mean of the negative dyadic logarithms of the probability $\bar{p}$ per symbol.

Similarly, the average amount of information per symbol of the original series, if these symbols have unequal relative frequencies or probabilities, is given by the arithmetic mean of the negative dyadic logarithms of the probabilities per symbol:

$$H = - \Sigma p_i \operatorname{ld} p_i \tag{4}$$

Their quotient

$$h = \frac{H}{H'} = \frac{\Sigma p_i \operatorname{ld} p_i}{\operatorname{ld} \bar{p}} \tag{5}$$

shows the influence of weighting the symbol ranks by the actual probabilities, or roughly, the influence of taking the statistical distribution of the symbols into account.

Because of their formal similarity with the expression for the *Entropy* of statistical mechanics, H and H' are called Entropies, and their quotient h the relative entropy. Its complement

$$R = 1 - h \tag{6}$$

is called the Redundancy of the code.

It is the property of a coding system which enables us to make guesses, with a reasonable degree of expectation to be correct, as to missing parts of a message. The quantity

$$R = 1 - h$$

enables us to compare percentage series with different numbers of classes.

Applying these ideas now to the protein mixture, we interpret H as measuring the number of guesses needed for identifying a particle (molecule) in a given complete mix. The mixture of protein constituents—or any other mixture belonging to living matter—is not stable, that is, its particles do not stay put, and its constituents must therefore be thought of as constantly changing place in a random way. H may therefore be also interpreted as the number of guesses needed for determining the nature of a particle in a given small spatial area, or, what is the same, the time needed for a specified particle to appear in a given small spatial area. All these formulations express one and the same concept: *the internal mobility of the mix*. The greater H, the smaller that mobility. We have therefore in the entropy an objective measure of that important quality of protein mixture: the internal mobility. This is not to be confused with viscosity which is the mobility relative to the environment.

Considering what has been said above, it would seem advantageous to use the entropy as a criterion for the therapeutic comparison. The following table gives the distribution of serum constituents in two samples of normal serum, and in five types of pathological serum.

Table 2

	Normal	Normal	Cirrhosis	Plasmacytoma.	Nephrosis	Virus hepatitis	Lupus erythematodes
Albumin	61.3 %	60.1 %	29.3 %	35.6 %	33.7 %	36.3 %	27.4 %
Globulin α_1	4.1 %	4.4 %	4.1 %	3.2 %	5.7 %	4.4 %	10.1 %
Globulin α_2	8.1 %	11.9 %	6.9 %	8.8 %	27.2 %	3.7 %	21.5 %
Globulin β	11.0 %	13.3 %	14.7 %	11.0 %	21.4 %	15.1 %	17.6 %
Globulin γ	15.5 %	10.3 %	45.0 %	41.4 %	21.0 %	40.5 %	23.4 %
	H= 1.683	H= 1.738	H= 1.900	H= 1.874	H= 2.224	H= 1.845	H= 2.256

The comparison of the entropies in the above table shows:

(1) That the entropies of the two normal sera are very similar.

(2) That the entropies for the degenerative diseases of nephrosis and cirrhosis of the liver are characteristically different from the entropies for the normal serum.

(3) That the entropy for the dermatological disease (lupus erythematodes) almost reaches the maximum possible entropy which, in this case, is $H' = 2.322$ (see formula 3).

(4) The entropy for plasmacytoma and virus hepatitis, although different from that for normal serum does not show a very pronounced difference.

What should be kept in mind is that although the entropy appears to characterise changes in serum composition in a very effective way, yet for an obvious reason can not be taken as a 100% substitute for considering the percentage series in detail. According to the calculation of the entropy, this quantity does not take into account the order or rank in which the different percentages appear. If, therefore, in a pathological serum the change would consist in the percentages for two constituents merely changing place, the entropy would remain the same in spite of the curve and the percentage series becoming very different. Although this is not a case which will happen very often, yet the possibility exists and makes it advisable not to forget the curve and the whole series altogether when using the entropy as a therapeutic criterion.

It was said above that the entropy does not take into account the mere change-over of percentages among the five constituents. In a purely formal way this is true. That is, the quantity of H would not change. However, this does not mean that the whole method of information-theory is insensitive to such a change-over. Mere changing of the proportions among the five constituents, although not altering H, admits a very interesting conclusion about the nature of the mix. Looking at the composition of the normal serum, it appears that the greatest proportion is contributed by the lightest constituent (albumin). In the light of the interpretation given above of H obtained for such mixtures, this means that the particle which most frequently changes place in the mix would be the lightest. If we now find that in a disease like nephrosis the proportion of the heavier particles, and in cirrhosis that of the heaviest particles, namely gamma globulin, has considerably increased, it means that it is the heavier or the heaviest particle which accounts for most of the movement in the mix. The idea of using the entropy as a nosographic criterion in certain diseases is therefore twofold. *First*, the greater the entropy, the less the internal mobility of the mix. *Secondly*, the greater the deviation of the distribution from that of a normal serum, and especially the greater the proportions of the heavier components, the more abnormal will be the physical characteristics of that mobility. Instead of the lighter particles accounting for most of the movement in the mix, it will then be the

heavier particles which do so. As a rule, this would result in a slowing down of the internal movement.

The entropy difference may be tested for significance as follows. Having a number of determinations for normal sera and therefore a number of entropies for normal sera, we can calculate the standard deviation of the entropy in a normal serum. The entropy then of a pathological serum in which we are interested may be compared with the average entropy for the normal serum (representing the "population") using a test of statistical significance.

Example. The variability of the entropy of normal sera will be a function of the experimental errors, biological variations, systematically different methods of experiment and errors in evaluation. In what follows, the standard deviation is calculated so as to take the variability due to all these factors into account. It will therefore be somewhat greater than the standard deviation calculated by taking only one or the other of these factors into account.

Table 3. Serum Proteins of Normal Sera.

Proteins	1	2	3	4	5	6	7	8	9	10
Albumin	61.3	62.5	60.1	60.1	59.0	59.0	62.8	65.1	63.4	63.0
Globulin α_1	4.1	4.7	3.5	4.4	4.3	5.6	5.8	5.4	5.7	5.8
Globulin α_2	8.1	8.8	7.4	11.9	7.7	9.2	6.5	6.1	6.6	6.8
Globulin β	11.0	11.8	7.3	13.3	12.5	13.5	8.3	7.8	8.4	8.2
Globulin γ	15.5	16.7	14.3	10.3	16.5	12.6	16.6	15.6	15 9	16.2
Entropy H	1.683	1.764	1.628	1.731	1.734	1.764	1.644	1.584	1.634	1.644

Series (1) was obtained by *Grassmann* [1952] as the average of 25 normal sera.

Series (2) and (3) were obtained by adding and subtracting respectively, the standard deviation for each of the protein components of series (1) and thus taking into account the biological variations.

Series (4) is also due to *Grassmann* [1951], but was obtained from other sera than series (1).

Series (5) was obtained by *Broicher* and *Odenthal* [1954] from a normal serum.

Series (6) is an average of 5 electrophoresis determinations by the *Tiselius* method (*Grassmann* [1952]).

Series (7–10) (*Grassmann* [1954]) represent each an average of 5 determinations of one and the same serum, but they were obtained by 4 different methods of evaluation of paper electrophoresis data.

The average of the 10 values of the entropy results as:

$$H = 1.681$$

The standard deviation results as:

$$\sigma = \sqrt{\frac{28.294 - 10 \times 1.681^2}{9}} = 0.064$$

Twice the standard deviation added to and subtracted from the average H gives the two limiting values of 1.809 and 1.553 respectively. A value of more than 1.809 or less than 1.553 could be regarded as abnormal.

Returning now to Table 2, we find that the *entropies calculated from non-normal sera are all beyond the two standard error limit for normal sera.* They all exceed the average for normal sera by more than twice the standard deviation. This shows the results of the statistical significance test to be in full agreement with the clinical findings according to which the persons from whom the sera of Table 2, col. 3–10 were obtained represented non-normal cases.

It thus shows the suitability of the proposed parameter, viz., the entropy of the serum mixture as a therapeutic criterion. In this example we have compared abnormal sera with normal sera. In quite an analogous way, we would compare the entropy for serum proteins obtained from persons suffering from a given disease in both the control and the experimental group.

Summary

After discussing certain shortcomings of conventional statistical significance tests when applied to electrophoresis data, a new method is suggested which consists in comparing the "Entropy" of serum protein mixtures calculated as the negative average logarithmic probability of the serum constituents.

It is shown that this parameter is an index of the internal mobility of the mixture of proteins, and of the preponderance of abnormal constituents in it, furthermore that its value obtained from clinically ascertained abnormal sera is such as would result for normal sera less often than 5 times in 100, considering the

variations of normal sera due to experimental, graphical, and observational errors. This shows the Entropy to be a suitable statistical criterion for distinguishing between normality and abnormality of electrophoresis data.

Résumé

Après avoir discuté certaines insuffisances des tests statistiques conventionnels si on les applique aux données de l'électrophorèse, l'auteur suggère une nouvelle méthode consistant à comparer l'«entropie» de mélanges de protéines sériques calculés comme la probabilité logarithmique moyenne des constituants du sérum.

Il est démontré que ce paramètre est un index de la mobilité interne du mélange de protéines et de la prépondérance de certains constituants anormaux; plus encore, que les résultats obtenus à partir de sérums confirmés cliniquement comme anormaux, ne sont comparables à ceux obtenus pour des sérums normaux que moins de 5 fois sur 100, si l'on tient compte des erreurs liées à l'observation, à l'expérimentation et à la transcription graphique. Cela montre que l'«entropie» est un critère statistique satisfaisant, permettant de distinguer entre normal et anormal dans les données de l'électrophorèse.

Zusammenfassung

Nach der Besprechung gewisser Mängel konventioneller statistischer Signifikanzproben, angewendet bei Daten der Elektrophorese, wird eine neue Methode vorgeschlagen, welche darin besteht, daß man die «Entropie» von Serumproteinmischungen vergleicht, errechnet als die durchschnittliche, logarithmische Wahrscheinlichkeit der Serumbestandteile.

Es wird aufgezeigt, daß dieser Parameter ein Index der inneren Beweglichkeit der Mischung der Proteine und des Übergewichtes der darin enthaltenen, abnormen Bestandteile ist, fernerhin, daß sein Wert, der durch klinische Feststellung abnormer Sera errechnet wird, ein solcher ist, wie er für normale Sera in weniger als fünf Fällen von hundert resultieren würde, wenn man die Veränderungen von normalen Sera mit in Erwägung zieht, die auf experimentelle, graphische und Beobachtungsfehler zurückzuführen sind. Die Entropie erweist sich hiermit als geeignetes statistisches Kriterium für die Unterscheidung von Normalität und Abnormität von Daten der Elektrophorese.

BIBLIOGRAPHY

1. *Broicher, H.* and *H. Odenthal:* Klin. Wschr. *32*, 592, 1954. – 2. *Grassmann, W.* and *K. Hannig:* Hoppe-Seyler's Ztschr. f. physiolog. Chem. *290*, 1, 1952. – 3. *Grassmann, W., K. Hannig* and *M. Knedel:* Dtsch. med. Wschr. *76*, 333, 1951. – 4. *Grassmann, W.:* Klin. Wschr. *1954*, 32. – 5. *Meyer-Eppler, W.:* Naturwissenschaften *39*, 341, 1952. – 6. *Shannon, C.E.:* Bell Syst. tech. J. *27*, 379, 1948. – 7. *Tiselius, A.:* Diss. Uppsala. 1930. Nova Acta Reg. Soc. Sci. upsal. (IV) 7, 4, 1930; Trans. Faraday Soc. *33*, 524, 1937. – 8. *Walther, A.:* Z. Ver. dtsch. Ing. *96*, 137, 1954.

Hogben, L.: Acta genet. *5*, 115–133, 1955

FIGURATE SERIES AND FACTORIAL NOTATION

By LANCELOT HOGBEN, Birmingham

1. Introduction

What follows explores an operative notation for the treatment of series of integers amenable to figurate representation. It will first be necessary to specify the symbols employed in the preliminary definitions. What we shall signify by the rank (r) and dimension (d) of a figurate series will be clear enough, if we set down the familiar Pythagorean family later specified as $F_{r,d}(1)$ for $r = 1, 2, \ldots 5$ and $d = 0, 1, 2, \ldots 6$:

d	RANK (r)					
	1	2	3	4	5	...
0	1	1	1	1	1	...
1	1	2	3	4	5	...
2	1	3	6	10	15	...
3	1	4	10	20	35	...
4	1	5	15	35	70	...
5	1	6	21	56	126	...
6	1	7	28	84	210	...

Here we can represent the terms of dimension 0 as points, those of dimension 1 as chains of r points, those of 2 as triangles whose sides are chains of r points, and those of dimension 3 as tetrahedra whose edges consist of r points also.

To define figurate series in the most general way, we shall employ the operation specified by

$$\nabla u_x \equiv u_x - u_{x-1} \tag{1}$$

We may then say that a figurate series of order $(n+1)$ is completely definable for all dimensions, if ∇ w.r.t. r:

$$\nabla F_{r\cdot d}(k_o, k_1 \cdots k_n) = F_{r\cdot(d-1)}(k_o, k_1 \cdots k_n) \tag{2}$$

$$F_{1\cdot o}(k_o, k_1 \cdots k_n) = k_o \tag{3}$$

$$F_{1\cdot 1}(k_0, k_1 \cdots k_n) = k_1 \tag{4}$$

$$F_{1\cdot 2}(k_0, k_1 \cdots k_n) = k_2 \tag{5}$$

. . . . . .

$$F_{1\cdot d}(k_0, k_1 \cdots k_n) = k_n \,(d > n) \tag{6}$$

For the specification of a formula, it will suffice to write $F_{r\cdot d}(c)$ if $k_x = c$ for all values of x, $F_{r\cdot d}(c,h)$ if $k_0 = c$ and $k_d = h$ $(d \geqslant 1)$, $F_{r\cdot d}(c,h,k)$ if $k_0 = c$, $k_1 = h$, $k_2 = k$ $(d \geqslant 2)$ *etc.* The accompanying table exhibits $F_{r\cdot d}(c,h,k)$ in the domain both of positive and of negative integers for the range $r = -4$ to $+4$ and for $d = 0$ to 4 inclusive. To express the general term in the same form for the domain of positive and negative integers alike it will be useful to define by analogy with $x^{(n)}$:

$$x^{[n]} = x\,(x+1)\,(x+2)\cdots(x+n-1) \tag{7}$$

In accordance with *Aitken's* usage of $x_{(n)} = \frac{x^{(n)}}{n!}$ we may then write

$$x_{[n]} = \frac{x^{[n]}}{n!} \text{ and } (-x)_{[n]} = (-1)^n x_{(n)} \tag{8}$$

In this notation:

$$x^{(n)} = (x-n+1)^{[n]} \text{ and } x^{[n]} = (x+n-1)^{(n)} \tag{9}$$

We shall likewise postulate

$$x_{[0]} = 1$$

The morphological implication of the term *dimension* (fig. 1) in this context is clear enough if we confine ourselves to the Pythagorean figurates for which $k_d = 1$ for all values of d; but it requires scrutiny when $k_1 = 0$. Thus the terms of the series $F_{r\cdot d}(s,0,1)$ are amenable to representation (fig. 5) as s-sided regular polygonates both for $d = 1$ and for $d = 2$. Accordingly, we may speak of the polygonate of dimension 1 as a plane *shell* and that of dimension 2 as *packed*, being uniquely determined for r, if we fix k_2, by successive addition to its predecessor of a plane shell of rank r. Thus we might designate k_2 as the *nuclear* constant definitive of the core, the law of growth in dimension 2 being determined uniquely by (2) and by the values assigned to k_0, k_1 and k_2.

If $k_3 = 1 = k_2$ and $k_1 = 0$ or 1 the 3-dimensional figurate will be a packed polyhedrate tapering from its base to a point pyramidwise. It will thus be clear that three constants suffice to define only a limited variety of 3-dimensional patterns. More generally, we may postulate that an n-dimensional series is amenable to representation as a polyhedral shell, in which event by fixing the nuclear constant k_n, we can generate for $d = (n+1)$ a figurate lattice of any corresponding form by superposition of concentric shells about a definitive core. For instance, $F_{r\cdot d}(12,-6,2,6)$ defines a family amenable in the positive domain to 2-dimensional representation by shell cubes whose faces are square figurates $F_{r\cdot 2}(2,1)$. In the third dimension the term of rank r is a cubical lattice of r shells each face of the outermost shells being a square of rank r.

2. *The Additive Property*

In what follows, we shall confine ourselves to series of order $n \leqslant 4$ and the figurate patterns they yield in the positive domain $r \geqslant 1$. For the evaluation of the terms of a plane figurate $F_{r\cdot d}(c,h,k)$

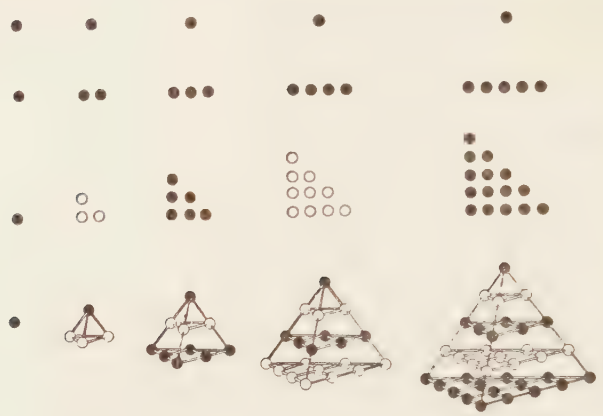


Fig. 1. The Family $F_{r\cdot d}(1)$, $d = 0$–3 and $r = 1$–5 inclusive.

of order 3, it suffices to define the constants c, h, k alone and in virtue of the definition, as seen from Table I, we may then write

$$\mathrm{F_{r\cdot d}(c,h,k) = k\cdot r_{[d-2]} + h(r-1)_{[d-1]} + c(r-1)_{[d]}} \tag{10}$$

More generally we may write:

$$\mathrm{F_{r\cdot d}(k_0,k_1\cdots k_n) = k_n\cdot r_{[d-n]} + \sum_{x=0}^{x=n-1} k_x\,(r-1)_{[d-x]}} \tag{11}$$

The law of formation of figurate series entitles us to extend them into the domain of negative integers as pointed out by *Frankel* [1950] for the particular case $F_{r\cdot d}(1)$. and independently by myself[1] for the superfamilies $F_{r\cdot d}(\overline{s-2},1,1)$ and $F_{r\cdot d}(s,0,1)$; but reliance on the customary notation of factorials and factorial powers assigns no manifest correspondence to expressions respectively specifying $F_{r\cdot d}(\cdots)$, and $F_{-r\cdot d}(\cdots)$, since

$$(r-1)_{[d]} = (r+d-2)_{(d)} \text{ and } (-r-1)_{[d]} = (-1)^d\,(r+1)_{(d)}$$

It is therefore pertinent to emphasise that (11) is equally valid for r positive or negative in the form cited. A sequence of numbers of receding rank in the negative domain will be amenable to figurate representation but will not necessarily tally in reverse order with positive terms in the same dimension, as is true of $F_{r\cdot d}(c.0.1)$ but of $F_{r\cdot d}(c,1)$ only if $-3 < c < 3$.

[1] *Frankel, E. T.* [1950]. A Calculus of Figurate Numbers and Finite Differences. *Amer. Math. Month.* LVII. – *Hogben, L.* Chance and Choice (London and New York).

Table I

r \ d	−4	−3	−2	−1	0	1	2	3	4
	c	c	c	c	c	c	c	c	c
0									
1			h−3c	h−2c	h−c	h	h+c	h+2c	h+3c
2		h−4c	k−3h+3c	k−2h+c	k−h	k	k+h+c	k+2h+3c	k+3h+6c
3	h−5c	k−4h+6c	−2k+3h−c	−k+h	0	k	2k+h+c	3k+3h+4c	4k+6h+10c
4	k−5h+10c	−3k+6h−4c	k−h	0	0	k	3k+h+c	6k+4h+5c	10k+10h+15c
	−4k+10h−10c	3k−4h+c							
	6k−10h+5c								

The advantage of exhibiting (10) as above without collecting terms will be evident, if we set $d = 2$ and $d = 3$, *viz.:*

$$F_{r\cdot 2}(c,h,k) = k + h(r-1) + c(r-1)_{[2]}$$
$$F_{r\cdot 3}(c,h,k) = k\cdot r + h(r-1)_{[2]} + c(r-1)_{[3]}$$

In these expressions r and $(r-1)$ are respectively chains of r and $(r-1)$ points, $(r-1)_{[2]}$ is a Pythagorean triangular number of rank $(r-1)$ and $(r-1)_{[3]}$ is a Pythagorean tetrahedral number of rank $(r-1)$. Given the formula for a figurate, we can thus see what patterns are possible in dimensions 2 and 3 by combining k points or lines length r, h lines or laminae of rank $(r-1)$ and c Pythagorean triangulates or Pythagorean tetrahedrates of rank $(r-1)$. The figurate pattern need not be unique as examples subsequently cited will sufficiently indicate.

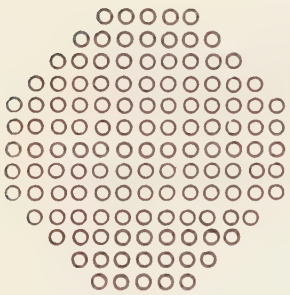
Conversely, the pattern leads to the formula by a fundamental *additive* property inherent in the notation employed, *viz:*

$$F_{r\cdot d}(a,b,c) \pm F_{r\cdot d}(u,v,w) = F_{r\cdot d}(a\pm u,\ b\pm v,\ c\pm w) \quad (12)$$

By the same token we may thus write

$$\pm K\cdot F_{r\cdot d}(u,v,w) = F_{r\cdot d}(\pm Ku,\ \pm Kv,\ \pm Kw)$$

We may thus employ either or both of two methods for the structural analysis of the figurate pattern: (a) resolution in the same dimension into standard elements, i.e. of the



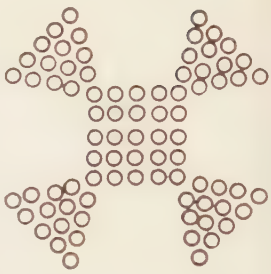


Fig. 2. $F_{r.2}$ (14, —3, 1) $F_{r.2}$ (12, 0, 1) $F_{r.2}$ (6, 5, 1)

plane figurate into triangulates usually of lower rank and of the solid figurate into tetrahedrates *ditto;* (b) resolution without loss of rank by elimination of elements of lower dimension, e.g. plane faces in the case of the solid, lines or points in the case of the plane figurate. Since (12) holds good only for the composition of elements of identical rank and identical dimension, its utilisation will invoke one or both of two operations which the 3rd order figurate will suffice to illustrate: (i) *The Dimensional Shift*

$$F_{r.(d-1)}(c,h,k) = F_{r\cdot d}(0,c,h,k)$$
$$F_{r\cdot(d-2)}(c,h,k) = F_{r\cdot d}(0,0,c,h,k)$$
$$F_{r\cdot(d-3)}(c,h,k) = F_{r\cdot d}(0,0,0,c,h,k)$$

(ii) *The Rank Shift*

$$F_{(r-1)\cdot d}(c,h,k) = F_{r\cdot d}(c,\ h-c,\ k-h,\ 0)$$
$$F_{(r-2)\cdot d}(c,h,k) = F_{r\cdot d}(c,\ h-2c,\ k-2h+c,\ -k+h,\ 0)$$
$$F_{(r-3)\cdot d}(c,h,k) = F_{r\cdot d}(c,\ h-3c,\ k-3h+3c,\ -2k+3h-c,\ k-h,\ 0)$$

Two examples invoking these operations will suffice to show how the figurate formula emerges from the figurate pattern by recourse to (11).

(a) The *atypical octagonate* whose formula is $F_{r.2}(14,-3,1)$ defines in the positive domain the series: 1, 12, 37, 76, 129 *etc.* On inspection (fig. 2) we see that it resolves into: (a) a cross consisting of 5 Pythagorean quadrates of rank r, i.e. $F_{r.2}(2.1)$, with elimination of 4 chains of r points; (b) four Pythagorean triangles $F_{(r-2)\cdot 2}(1)$ of rank $(r-2)$. Thus the build-up is:

$$\begin{aligned} &5F_{r\cdot 2}(2,1) - 4F_{r\cdot 1}(1) + 4F_{(r-2)\cdot 2}(1) \\ &\quad = 5F_{r\cdot 2}(2,1) - 4F_{r\cdot 2}(0,1) + 4F_{r\cdot 2}(1,-1,0) \\ &\quad = F_{r\cdot 2}(10,5,5) + F_{r\cdot 2}(0,-4,-4) + F_{r\cdot 2}(4,-4,0) \\ &\quad = F_{r\cdot 2}(14,-3,1) \end{aligned}$$

(b) The *packed octahedrate* whose formula is $F_{r\cdot 3}(4,0,1)$ defines in the positive domain the series: 1, 6, 19, 44, 85, 146 *etc.* We may look upon it as two apical pyramidates (*vide infra*) on a square base of rank r adhering base to base with elimination of a Pythagorean quadrate of the same rank. We may also envisage it as an apical pyramidate on a square base of rank r adhering by its base to another of rank $(r-1)$. Thus we may proceed as follows:

$$2F_{r\cdot 3}(2,1) - F_{r\cdot 2}(2,1) = F_{r\cdot 3}(4,2,2) + F_{r\cdot 3}(0,-2,-1) = F_{r\cdot 3}(4,0,1)$$

Alternatively,

$$F_{r\cdot 3}(2,1) + F_{(r-1)\cdot 3}(2,1) = F_{r\cdot 3}(2,1,1) + F_{r\cdot 3}(2,-1,0) = F_{r\cdot 3}(4,0,1)$$

3. Figurate Patterns

Before proceeding further it will be advantageous to introduce some descriptive terms to sidestep undue periphrasis. *Plane Figurates* are rectilinear 2-dimensional closed patterns of s sides each of r points including the 2 vertices if the rank is r.

Solid Figurates are rectilinear 3-dimensional patterns of V vertices, $\varnothing$ faces of s *sides* and with E edges, each edge containing r points including the 2 vertices if the rank is r.

Polygonates are plane figurates with an equal number of vertices and sides.

Plane Shells are polygonates whose points lie only on the sides, representable algebraically by a 1-dimensional formula, *viz:* $F_{r\cdot 1}(\ldots) = (r-1)\,s$.

Shell Polyhedra are solid figurates whose $\varnothing$ faces are plane shells representable algebraically by a 2-dimensional formula, *viz:* $F_{r\cdot 2}(\ldots) = V + (r-1)\,E$.

Regular Polyhedrates conform to the relations $E = \frac{1}{2}\varnothing s$ and $V + \varnothing - E = 2$.

Pyramidates are packed solid figurates the base of which is an s-sided polygonate and whose remaining sides are 3-sided polygonates, so that $\varnothing = (s+1) = V$ and $E = 2s$.

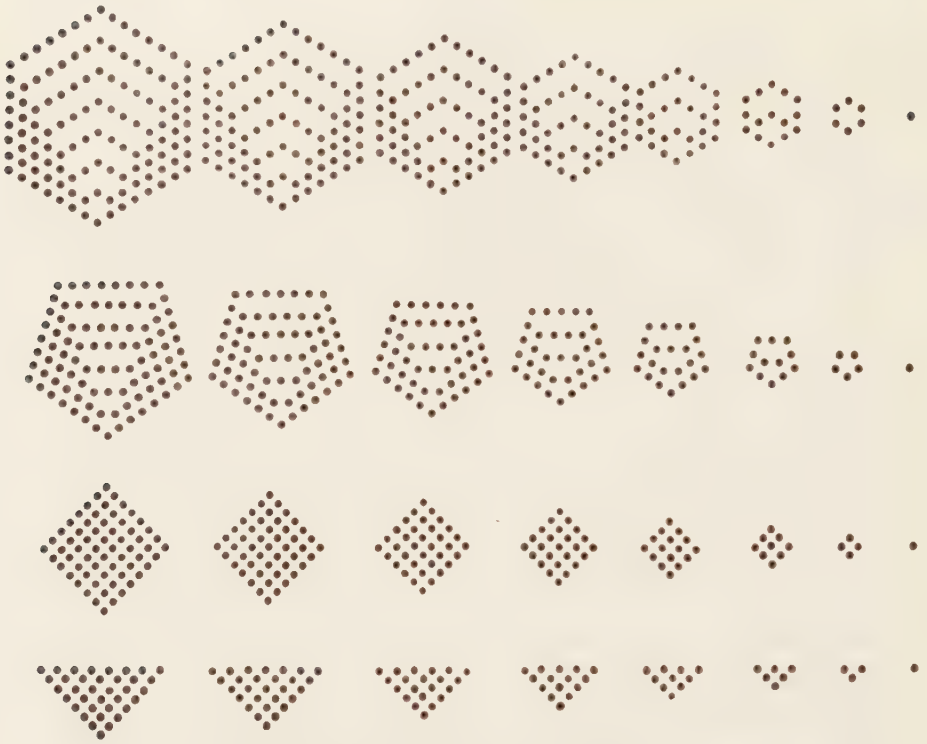


Fig. 3. The Family $F_{r\cdot 2}(\overline{s-2},\ 1)$ for $s = 3$–6 and $r = 1$–8 inclusive.

Prismates of rank r are formed by superposition of r polygonates of rank r and s sides, so that $V = 2s$, $E = 3s$ and $\varnothing = (s+2)$.

Latticeates are polyhedrates formed by concentric addition of shell polyhedrates.

Packed polyhedrates are not amenable to shell resolution.

4. *Plane Figurates*

Aside from anomalous types, such as the octagonate $F_{r\cdot 2}(14,\ 3,1)$ already mentioned, 2 classes of polygonates deserve special attention (figs. 3–6):

(a) *Apical Polygonates* $F_{r\cdot 2}(s-2,1)$

(b) *Central Polygonates* $F_{r\cdot 2}(s,0,1)$

Of these, the apical polygonates which grow (figs. 3 and 4) from a vertex include as a special case $F_{r\cdot 2}(1)$ which represents the sum

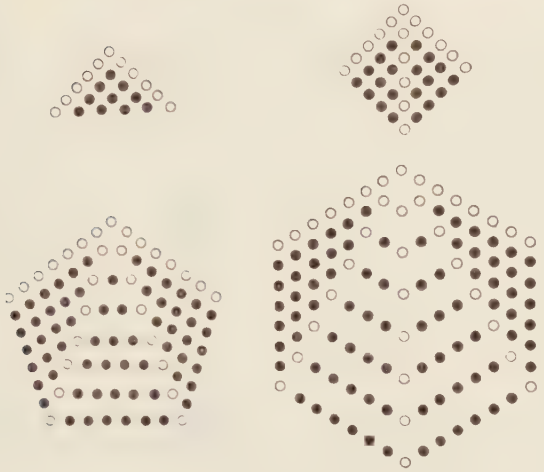


Fig. 4. Apical Growth of the Figurate Pattern $F_{r.2}(\overline{s-2}, 1)$

of the r natural numbers in the positive domain, being the 2-dimensional member of the family $F_{r.2}(1)$ whose diagonal terms are rows of Pascal's triangle. $F_{r.2}(2,1)$, $F_{r.2}(3,1)$ *etc.* represent quadrates, pentagonates, *etc.* $F_{r.2}(2,1)$ represents the squares of the first r odd numbers in the positive domain.

The central polygonates (figs. 5 and 6) as defined above grow centrifugally by addition of successive 2-dimensional shells to the nuclear constant 1, the corresponding 1-dimensional series being definitive of the shell polygonates. In the positive domain, the octagonal member of the family, i.e. $F_{r.2}(8,0.1)$ represents the squares of the first r odd numbers.

As stated, such series admit of more than one visual pattern. Thus we may visualise $F_{r.2}(10,1)$ as

(a) an *apical polygonate* of 12 sides;

(b) an apical *stellate* of degree 6, being an apical hexagonate of rank r on each side of which lies the base of an apical triangulate of rank $(r-1)$;

Similarly we may visualise $F_{r.2}(20,0,1)$ as:

(a) a *central polygonate* of 20 sides;

(b) a *central cruceate* of degree 5, being a central pentagonate of rank r to each of the vertices of which by its own vertex attaches a central triangulate of rank r with elimination of 5 vertical points in all;

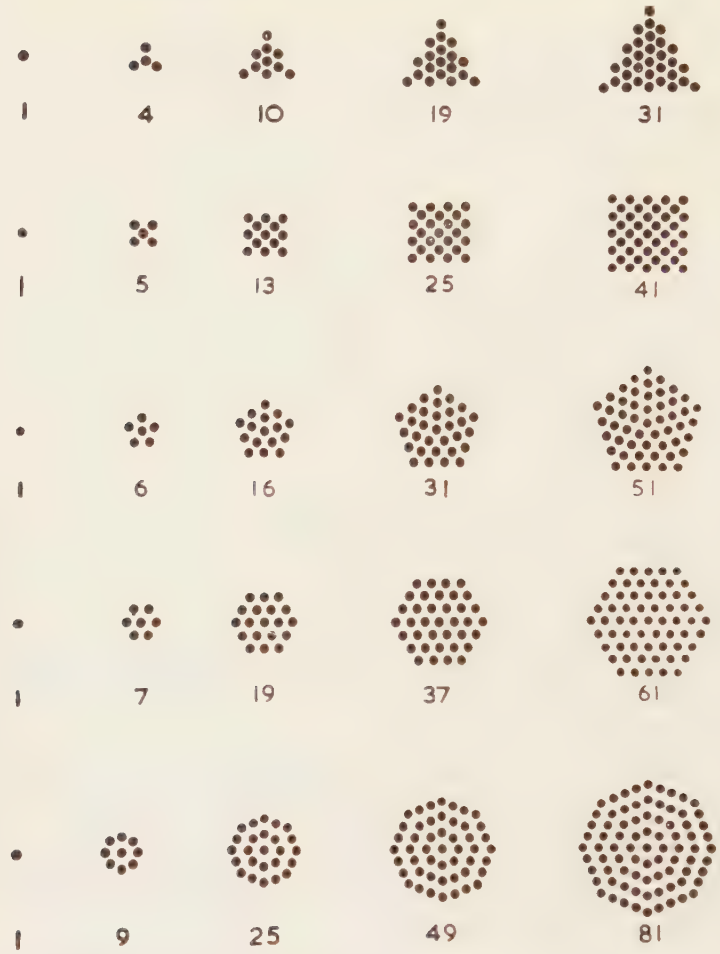


Fig. 5. The Family $F_{r.2}(s,0,1)$ for $s = 3$–6 inclusive and $s = 8$, $r = 1$–5 inclusive.

(c) a *central floreate* of degree 4, being a central quadrate to each of the vertices of which attaches a quadrate also of rank r with elimination of 4 vertical points in all.

Needless to say, the number of plane figurate patterns of order 2 or 3 is very great. Among many others one may cite, are the *apical cruceates* of degree p. The formula is $F_{r.2}(2p-2, p+1, 1)$. Each consists of p apical triangulates of rank r attached to the vertices of a p-sided apical polygonate also of rank r with elimination of r points.

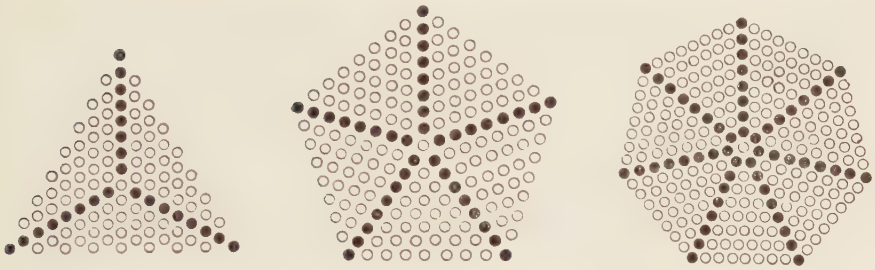


Fig. 6. Composition of the Central Polygonate $F_{r\cdot 2}(s, 0,1)$

5. *Solid Figurates*

These call for more detailed discussion in view of the possible utility of the notation proposed for descriptive purposes in connexion with the fine structure of matter. *Pyramidates* are 3-dimensional derivatives of plane figurates. Corresponding to the two classes of polygonates mentioned we have $F_{r\cdot 3}(s{-}2, 1)$ and $F_{r\cdot 3}(s, 0,1)$. The central pyramidate on a hexagonal base, i.e. $F_{r\cdot 3}(6,0,1)$ represents the cube of the rth natural number in the positive domain. The corresponding 4th dimensional series therefore leads at once to the formula for the sum of the cubes, whence we note that $F_{r\cdot 4}(6.0.1)$ is the square of $F_{r\cdot 2}(1)$. The central pyramidate of degree 8, i.e. $F_{r\cdot 3}(8,0,1)$ yields for $d = 4$ the sum of the squares of the first r odd numbers.

Prismates may be generalised by superposition of apical or central polygonates of the same rank. If s and $\varnothing$ have their usual meaning the formulae are:

Apical $F_{r\cdot 3}(3s{-}6, 4{-}s, 1)$ *or* $F_{r\cdot 3}(3\varnothing, 6{-}\varnothing, 1)$

Central $F_{r\cdot 3}(3s, -s, 1)$ *or* $F_{r\cdot 3}(3\varnothing{-}6, 2{-}\varnothing, 1)$

Latticeates. We may build up the formulae for three types of regular polyhedral latticates to which we may refer as *shell-faced*, *apical* and *central* according as the facet is a shell, apical or central polygonate. In addition to the edges, which it shares with contiguous facets, each facet of the apical shell will contain residual points defined by:

$$F_{r\cdot 2}(s{-}2, 1, 1) - F_{r\cdot 1}(s, 0) = F_{r\cdot 2}(s{-}2, 1{-}s, 1)$$

Likewise each facet of the central type will contain residual points defined by:

$$F_{r\cdot 2}(s, 0, 1) - F_{r\cdot 1}(s, 0) = F_{r\cdot 2}(s, -s, 1)$$

For the shell-faced shell with E edges, V vertices and $\varnothing$ faces the formula is:

$$F_{r\cdot 1}(E, V-E, 0) = F_{r\cdot 2}(0, \tfrac{1}{2}\varnothing s, 2-\varnothing, 0)$$

Thus the apical shell is:

$$F_{r\cdot 2}(0, \tfrac{1}{2}\varnothing s, 2-\varnothing, 0) + \varnothing F_{r\cdot 2}(s-2, 1-s, 1, 1)$$
$$= F_{r\cdot 2}(\varnothing s - 2\varnothing, \varnothing - \tfrac{1}{2}\varnothing s, 2, \varnothing)$$

Similarly, the central type of shell is:

$$F_{r\cdot 2}(0, \tfrac{1}{2}\varnothing s, 2-\varnothing, 0) + \varnothing F_{r\cdot 2}(s, -s, 1, 1)$$
$$= F_{r\cdot 2}(\varnothing s, -\tfrac{1}{2}\varnothing s, 2, \varnothing)$$

The latticates are thus:

Shell	*Apical*	*Central*
Tetrahedrate		
$F_{r\cdot 2}(6, -2, 0)$	$F_{r\cdot 3}(4, -2, 2, 4)$	$F_{r\cdot 3}(12, -6, 2, 4)$
Cubate		
$F_{r\cdot 2}(12, -4, 0)$	$F_{r\cdot 3}(12, -6, 2, 6)$	$F_{r\cdot 3}(24, -12, 2, 6)$
Octahedrate		
$F_{r\cdot 2}(12, -6, 0)$	$F_{r\cdot 3}(8, -4, 2, 8)$	$F_{r\cdot 3}(24, -12, 2, 8)$
Dodecahedrate		
$F_{r\cdot 2}(30, -10, 0)$	$F_{r\cdot 3}(36, -18, 2, 12)$	$F_{r\cdot 3}(60, -30, 2, 12)$
Icosahedrate		
$F_{r\cdot 2}(30, -18, 0)$	$F_{r\cdot 3}(20, -10, 2, 20)$	$F_{r\cdot 3}(60, -30, 2, 20)$

In these formulae, the nuclear constant k_3 is arbitrary and we may accordingly step up or step down each term in the 3-dimensional series by assigning a value other than as here given.

Packed Regular Polyhedrates. We have already seen that we can construct a packed octahedrate by base to base juxtaposition of pyramidates on a quadrate base. If the pyramidates are of the apical type the formula is $F_{r\cdot 3}(4, 0, 1)$ which is alternatively amenable to visualisation as a central pyramidate on a 4-sided base. The central octahedrate formed in the same way is $F_{r\cdot 3}(8, -4, 2, 1)$. The cube of rank r formed by superposition of r apical quadrates is the apical prismate $F_{r\cdot 3}(6, 0, 1)$, alternatively a central hexagonal pyramidate. The cube of rank r formed by superposition of r central quadrates is $F_{r\cdot 3}(12, -4, 1)$. The packed octahedrates mentioned are particular examples of solid figurates which one might speak of as

bipyramidates, being formed by base to base juxtaposition of rank r pyramidate whose base has s sides and the corresponding pyramidate of rank $r\,(r-1)$. The number of faces is $2s$. The general formula for the apical type is $F_{r\cdot 3}(2s-4, 4-s, 1)$. For the central type, it is $F_{r\cdot 3}(2s, -s, 2, 1)$. Two members of the latter merit mention. $F_{r\cdot 4}(48, -24, 2, 1)$ defines the sum of the cubes of the first r odd numbers, and $F_{r\cdot 5}(24, -12, 2, 1)$ defines the sum of the 4th powers of the first r natural numbers.

Among other classes of packed polyhedrates which one might mention are solid stellates which are figures formed by base to face fusion of pyramidates with a polyhedrate core of rank r and elimination of the 2-dimensional base of each pyramidate. If the core is a packed cube four cases arise:

(a) Core and pyramidates apical:

$$\begin{aligned} &F_{r\cdot 3}(6.0.1) + 6F_{(r-1)\cdot 3}(2,1) \\ &\quad = F_{r\cdot 3}(6.0.1) + 6F_{r\cdot 3}(2,-1,0) \\ &\quad = F_{r\cdot 3}(18,-6,1) \end{aligned}$$

(b) Core apical, pyramidates central:

$$\begin{aligned} &F_{r\cdot 3}(6,0,1) + 6F_{(r-1)\cdot 3}(4,0,1) \\ &\quad = F_{r\cdot 3}(6,0,1) + 6F_{r\cdot 3}(4,-4,1,0) \\ &\quad = F_{r\cdot 3}(30,-24,7,1) \end{aligned}$$

(c) Core central, pyramidates apical:

$$\begin{aligned} &F_{r\cdot 3}(12,-4,1) + 6F_{r\cdot 3}(2,1) \\ &\quad = F_{r\cdot 3}(12,-4,1) + 6F_{r\cdot 3}(2,-1,0) \\ &\quad = F_{r\cdot 3}(24,-10,1) \end{aligned}$$

(d) Core central, pyramidates central:

$$\begin{aligned} &F_{r\cdot 3}(12,-4,1) + 6F_{(r-1)\cdot 3}(4,0,1) \\ &\quad = F_{r\cdot 3}(12,-4,1) + 6F_{r\cdot 3}(4,-4,1,0) \\ &\quad = F_{r\cdot 3}(36,-28,7.1) \end{aligned}$$

6. *Figurate Series and Binomial Coefficients in the Negative Domain*

If we extend (Table II *below*) the terms of $F_{r\cdot d}(1)$ in the domain of negative integers, we at once notice that the terms of the column headed by $F_{r\cdot 0}$ are those of the expansion of $(1-r)^r$ followed by an infinitude of zeros. Diagonal terms running leftwards and downwards from $F_{r\cdot d}$ terminate in the same way and successively for those beginning with $F_{1\cdot 0}(1)$, $F_{2\cdot 0}(1)$, $F_{3\cdot 0}(1)$ *etc.* generate successive terms of $(1+1)^0$, $(1+1)^1$, $(1+1)^2$ *etc.* The diagonal series which start with $F_{-r\cdot 0}$ running leftwards and downwards do not terminate.

Table II

r	–6	–5	–4	–3	–2	–1	0	1	2	3	4	5	6	...	d
...	1	1	1	1	1	1	1	1	1	1	1	1	1	...	0
...	–6	5	–4	–3	–2	–1	0	1	2	3	4	5	6	...	1
...	15	10	6	3	1	0	0	1	3	6	10	15	21	...	2
...	–20	–10	–4	–1	0	0	0	1	4	10	20	35	56	...	3
...	15	5	1	0	0	0	0	1	5	15	35	70	126	...	4
...	–6	1	0	0	0	0	0	1	6	21	56	126	252	...	5
...	1	0	0	0	0	0	0	1	7	28	84	210	462	...	6
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

They yield successive terms for the expansion of a binomial with an index which is a negative integer.

The law of formation of such binomial coefficients is consistent alike in the domain of positive and negative values of r for $F_{r\cdot d}(1)$. If $B_{r\cdot n}$ and $B_{r\cdot -n}$ respectively denote the term of rank r in the expansion of $(1+1)^n$ and $(1+1)^{-n}$, we see by inspection of the diagonal series that

$$B_{r\cdot n} = F_{(n-r+1)\cdot r}(1) \textit{ and } B_{r\cdot -n} = F_{(-n-r+1)\cdot r}(1) \qquad (13)$$

Whence from (10):

$$B_{r\cdot n} = (n-r+1)_{[r]} = n_{(r)} = \frac{n!}{(n-r)!}$$

$$B_{r\cdot -n} = (-n-r+1)_{[r]} = (-n)_{(r)} = (-1)^r \cdot n_{[r]} = \frac{(-1)^r \cdot (n+r-1)!}{(n-1)!}$$

Needless to say, we can define in both domains a shadow family of diagonates for any family of figurates, including *Pascal*'s numbers as a special case $B_{r\cdot n} = P_{r\cdot n}(1)$, if we write more generally:

$$P_{r\cdot n}(k_0, k_1 \ldots k_n) = F_{(n-r+1)\cdot n}(k_0, k_1 \ldots k_n) \qquad (14)$$

In the positive domain, $P_{r\cdot n}(s\ 2,\ 1)$ derived from the superfamily whose 2-dimensional representation is a regular apical polygonate specifies successive terms of $(1-s-2)(1+1)^{n-1}$.

7. *Vandermonde Expansions*

We may analyse a figurate pattern by breaking it down into figurates of lower rank than $(r\ 1)$. Thus the make-up of the central *hexagonate* $F_{r\cdot 2}(6, 0, 1)$ is the composition of a point $F_{r\cdot 0}(1)$, 6 chains

of $(r-1)$ points, i.e. $6F_{(r-1)\cdot 1}(1)$ and 6 Pythagorean triangulates of rank $(r-2)$, i.e. $6F_{(r-2)\cdot 2}(1)$:

$$\begin{aligned} &F_{r\cdot 0}(1) + 6F_{(r-1\cdot)1}(1) + 6F_{(r-2)\cdot 2}(1) \\ &\quad = F_{r\cdot 2}(0, 0, 1) + F_{r\cdot 2}(0, 6, 0) + F_{r\cdot 2}(6, -6, 0) \\ &\quad = F_{r\cdot 2}(6, 0, 1) \end{aligned}$$

In performing this operation we implicitly pass into the domain of negative integers without violating (12); and we have already seen that the admissibility of such an extension of figurate series leads to a consistent definition of binomial coefficients referable to exponents which are either positive or negative integers in virtue of the identities $(-x)^{(n)} = (-1)^n . x^{[n]}$ and $(-x)^{[n]} = (-1)^n . x^{(n)}$.
When n is negative[1]:

$$x^{(-n)} = \frac{1}{(x+1)^{[n]}} \text{ and } x^{[-n]} = \frac{1}{(x-1)^{(n)}} \tag{15}$$

Just as we may express u_n in terms of $\triangle^x u_0$ $(x = 0, 1, 2 \ldots n)$ by recourse to the operation $(1+\triangle)^x$, we may express u_{-n} in terms $\nabla^x u_0$ as defined by (1) by recourse to the operation $(1-\nabla)^x$. Whence we have:

$$x_r = \sum_{u=0}^{u=\infty} r_{(u)} \triangle^u x_0 \text{ and } x_{-r} = \sum_{u=0}^{u=\infty} (-1)^u r_{(u)} \nabla^u x_0 \tag{16}$$

Let us now consider the following series extending into the domain of both negative and positive integers

Rank (r)	$A_{(r)}$	$A_{[r]}$
b	$(a+b)^{(n)}$	$(a+b)^{[n]}$
b−1	$(a+b-1)^{(n)}$	$(a+b-1)^{[n]}$
---	-------	-------
2	$(a+2)^{(n)}$	$(a+2)^{[n]}$
1	$(a+1)^{(n)}$	$(a+1)^{[n]}$
0	$a^{(n)}$	$a^{[n]}$
−1	$(a-1)^{(n)}$	$(a-1)^{[n]}$
−2	$(a-2)^{(n)}$	$(a-2)^{[n]}$
---	-------	-------
−b+1	$(a-b+1)^{(n)}$	$(a-b+1)^{[n]}$
−b	$(a-b)^{(n)}$	$(a-b)^{[n]}$

[1] These definitions are in line with the usage of *Aitken*, and *Milne Thomson* in contradistinction to that of *Boole* and *Steffensen*. As *Freeman* points out, *Boole*'s definition, though (like the above) consistent with the formal identity of $\triangle x^{(r)}$ *etc.*, fails to satisfy the relations:

$$x^{(n)} = x^{(m)}(x-m)^{(n-m)} \text{ and } x^{[n]} = x^{[m]}(x+m)^{(n-m)}$$

Whence we may write (r being an integer):

$$A_{(r)} = (a+r)^{(n)}; \quad \triangle^u . A_{(0)} = n^{(u)} . a^{(n-u)}; \quad \nabla^u . A_{(0)} = n^{(u)} . (a-u)^{(n-u)}$$

$$(a+b)^{(n)} = \sum_{u=0}^{u=\infty} b_{(u)} . n^{(u)} . a^{(n-u)} = \sum_{u=0}^{u=\infty} n_{(u)} . b^{(u)} . a^{(n-u)} \qquad (17)$$

The term of rank $-b$ is likewise

$$(a-b)^{(n)} = \sum_{u=0}^{u=\infty} (-1)^u . n_{(u)} . b^{(u)} (a-u)^{(n-u)} \qquad (18)$$

Similarly, we may define a series:

$$A_{[r]} = (a+r)^{[n]}; \quad \triangle^u A_{[0]} = n^{(u)} (a+u)^{[n-u]}; \quad \nabla^u A_{[0]} = n^{(u)} . a^{[n-u]}$$

Whence we derive:

$$(a+b)^{[n]} = \sum_{u=0}^{u=\infty} n_{(u)} . b^{(u)} (a+u)^{[n-u]} \qquad (19)$$

$$(a-b)^{[n]} = \sum_{u=0}^{u=\infty} (-1)^u . n_{(u)} . b^{(u)} . a^{[n-u]} \qquad (20)$$

Equations (18) and (19) in the form given fall out of step; but it is easy be recourse to the identity $n_{(u)} = n_{(v)}$ if $v = (n-u)$ to shew that

$$\sum_{u=0}^{u=\infty} (-1)^u n_{(u)} b^{(u)} (a-u)^{(n-u)} = \sum_{u=0}^{u=\infty} (-1)^u b^{[u]} a^{(n-u)}$$

$$\sum_{u=0}^{u=\infty} n_{(u)} b^{(u)} (a+u)^{[n-u]} = \sum_{u=0}^{u=\infty} n_{(u)} b^{[u]} a^{[n-u]}$$

Thus *Vandermonde*'s Theorem defined for the expansion of $(a+b)^{(n)}$ is one of four comparable expansions

$$(a+b)^{(n)} = a^{(n)} + na^{(n-1)}b + n_{(2)}a^{(n-2)}b^{(2)} + n_{(3)}a^{(n-3)}b^{(3)} + n_{(4)}a^{(n-4)}b^{(4)} \ \ldots \ etc. \qquad (21)$$

$$(a-b)^{(n)} = a^{(n)} - na^{(n-1)}b + n_{(2)}a^{(n-2)}b^{[2]} - n_{(3)}a^{(n-3)}b^{[3]} + n_{(4)}a^{(n-4)}b^{[4]} \ \ldots \ etc. \qquad (22)$$

$$(a+b)^{[n]} = a^{[n]} + na^{[n-1]}b + n_{(2)}a^{[n-2]}b^{[2]} + n_{(3)}a^{[n-3]}b^{[3]} + n_{(4)}a^{[n-4]}b^{[4]} \ \ldots \ etc. \qquad (23)$$

$$(a-b)^{[n]} = a^{[n]} - na^{[n-1]}b + n_{(2)}a^{[n-2]}b^{[2]} - n_{(3)}a^{[n-3]}b^{[3]} + n_{(4)}a^{[n-4]}b^{[4]} \ \ldots \ etc. \qquad (24)$$

For negative values we may write

$$A_{(b)} = (a \pm b)^{(-n)} \text{ and } A_{[b]} = (a \pm b)^{[-n]}$$

$$\triangle^u A_{(0)} = (-n)^{(n)} . a^{(-n-u)} = \frac{(-1)^n . n^{[u]}}{(a+1)^{[n+u]}}$$

$$\nabla^u A_{(0)} = (-n)^{(u)} (a-u)^{(-n-u)} = \frac{(-1)^n . n^{[u]}}{(a+u-1)^{[n+u]}} = \frac{(-1)^n . n^{[u]}}{(a+u-1)^{[n+u]}}$$

$$\triangle^u A_{[0]} = (-n)^{(u)} (a+u)^{[-n-u]} = \frac{(-1)^n . n^{[u]}}{(a+u-1)^{[n+u]}} = \frac{(-1)^n . n^{[u]}}{(a-n)^{[n+u]}}$$

$$\nabla^u A_{[0]} = (-n)^{(u)} . a^{[-n-u]} = \frac{(-1)^n . n^{[u]}}{(a-n)^{[n+u]}}$$

Whence we derive

$$(a+b)^{(-n)} = \sum_{u=0}^{u=\infty} \frac{(-1)^n . n_{[u]} . b^{(u)}}{(a+1)^{[n+u]}} \tag{25}$$

$$(a-b)^{(-n)} = \sum_{u=0}^{u=\infty} \frac{n_{[u]} . b^{[u]}}{(a+n)^{[n+u]}} \tag{26}$$

$$(a+b)^{[-n]} = \sum_{u=0}^{u=\infty} \frac{(-1)^n . n_{[u]} . b^{(u)}}{(a-n)^{[n+u]}} \tag{27}$$

$$(a-b)^{[-n]} = \sum_{u=0}^{u=\infty} \frac{n_{[u]} . b^{[u]}}{(a-1)^{[n+u]}} \tag{28}$$

The series cited for $(a \pm b)^{(n)}$ and $(a \pm b)^{[n]}$ are valid for all rational values of a, b and integral positive values of n. If n is a negative integer, those given above are convergent by *Raabe*'s test and by that of *Gauss*, when $(a + b + 2) > 1$. Thus the condition of convergence, when n is a negative integer, is

$$(a+b) > -1$$

8. Figurate Pattern and Figurate Structure

A possible application of a 3-dimensional figurate calculus may emerge from the following considerations. Consider an ordered system of particles of two sorts P_a and P_b, mass respectively W_a and W_b arranged in a lattice of interlacing shells or layers so that we may assign to each composite shell or layer a specification involving the two components $F_{r.2}(k_{0.a} \ldots k_{n.a})$ and $F_{r.2}(k_{0.b} \ldots k_{n.b})$. The volume $(V_{r.d})$ of the system will depend on r alone, if we fix

$d = 3$. Accordingly, we may specify any physical property dependent on V_r alone by the number of particles *(v)* in a lattice V_r as

$$V_{r.3} = F_{r.3}\ (k_{0.a} + k_{0.b} \ldots . k_{n.a} + k_{n.b})$$

Any physical property which depends on the density (D_r) will however involve W_a and W_b. For a given shell or layer we may write

$$D_{r.2} = W_a . F_{r.2}(k_{0.a} \ldots k_{n.a}) + W_b . F_{r.2}(k_{0.b} \ldots k_{n.b})$$

Whence we obtain:

$$D_{r.3} = F_{r.3}(W_a . k_{o.a} + W_b . k_{0.b} \ldots . W_a . k_{n.a} + W_b . k_{n.b})$$

The figurate so specified is a pure construct which generates a series of integers which are proportional to the density of the lattice as we add additional shells or layers. Likewise, we may conceive a physical property $(E_{r.d})$ which depends only on volume and electrical charge (E_a, E_b) in which event E_a and E_b would replace W_a and W_b in the above.

Summary

Discrete series, such as the triangular and tetrahedral numbers, are amenable to figurate representation in 0–3 dimensions if reducible in terms of a general expression involving arbitrary constants definitive of successive dimensions and of the sides of a 2-dimensional representation. By means of two operations, the notation advanced makes it possible to resolve any 3-dimensional or 2-dimensional figurate so defined into elementary figurates of lower dimension, whence it is possible to express in algebraic form any regular pattern of points in space subject to specification by rank, e.g. the arrangement of atoms in a crystal. The notation employed also exhibits the usual statement of *Vandermonde*'s theorem as a special case of a more general relation.

Résumé

Certaines suites discrètes, tels les nombres triangulaires ou tétraédriques, sont susceptibles d'être représentées en configurations de 0 3 dimensions, si elles sont réductibles aux termes d'une expression générale contenant des constantes arbitraires définissant des dimensions successives et les côtés d'une représentation à 2 dimensions. Au moyen de deux opérations, la notation proposée permet la réduction de toute configuration à 3 ou 2 dimensions ainsi définie

en configurations élémentaires de dimensions moindres, d'où il résulte qu'il est possible d'exprimer, sous forme algébrique, n'importe quelle répartition régulière de points dans l'espace, compte tenu de la spécification par le rang. Tel est, par exemple, l'arrangement des atomes dans un cristal. Dans la notation employée, apparaît aussi, comme cas particulier d'une relation plus générale, la proposition habituelle du téorème de *Vandermonde*.

Zusammenfassung

Diskrete Reihen, wie trianguläre und tetraedrale Ziffern, sind brauchbar für die Darstellung in 0–3 Dimensionen, wenn sie in Glieder eines allgemeinen Ausdruckes verwandelt werden können, welcher beliebige Konstanten enthält, die definitiv sind für aufeinanderfolgende Dimensionen und für die Seiten einer zweidimensionalen Darstellung. Mit Hilfe von zwei Operationen macht es das entwickelte Zeichenschema möglich, jede drei- oder zweidimensionale Figur in elementare Figuren einer niedrigeren Dimension aufzulösen, wonach es möglich wird, jedes regelmäßige Schema von Punkten im Raum, welches der Spezifikation durch Linien bedarf, durch eine algebraische Formel auszudrücken, z.B. die Anordnung von Atomen in einem Kristall. Das verwendete Zeichenschema läßt auch die gebräuchliche Darstellung von *Vandermondes* Theorem als einen Spezialfall einer allgemeineren Beziehung erkennen.

Mannheimer, E.: Acta genet. *5*, 134—163

(From the State Institute for Human Genetics, Uppsala, Head: Prof. G. Dahlberg)

MORTALITY OF BREAST FED AND BOTTLE FED INFANTS A COMPARATIVE STUDY

By EDGAR MANNHEIMER, M.D., Stockholm

The statistical analysis was done at the State Institute for Human Genetics. To the head of the Institute, Professor *G. Dahlberg*, the author extends his sincere thanks for all his support and interest in the planning and carrying out of the study. The author also wishes to thank the Medical Officer of Stockholm, Med. Dr. *E. Rietz*, the Registrar at the Royal Central Office of Statistics, *Y. Fritzell*, the Registrars at the Statistical Office of the Town of Stockholm *O. Olinder* and *P.E. Anér*, and Mrs. *Irma Boström* for valuable assistance in the collecting of material.

The study was performed by means put at the disposal of AB Ferrosan, Malmö, Sweden.

From previous investigations it is well known that the prognosis for artificially fed infants is much worse than that for breast fed ones. This point of view dominated in ancient times also, which is fully shown in *Powers*' work, 1935.

The fundamental reports in this field are those performed at the suggestion and under the supervision of *R. Böckh* on material from Berlin. During a period of not less than 40 years (1878–1917) reports about the feeding of infants who had died were an integral part of Berlin's official statistics on the infant mortality rate. The material was divided into legitimate and illegitimate infants and, furthermore, into causes of death. A control material was available from the results of the population census in Berlin, taken regularly every fifth year. Likewise, on the initiative of *Böckh*, information about infant feeding was collected at the 1885–1910 census.

Thiemich and *Bessau* used the Berlin material and calculated the infant mortality rate of naturally and artificially fed infants on

three different occasions. Table 1 shows that there was throughout a considerably greater mortality rate in bottle fed infants, and also how this mortality decreased during the three decades in question. In 1885–86 there was in Berlin a 6.4 times greater risk of death for a bottle fed than for a breast fed infant, while in 1906 the corresponding figure was 3.7. The Berlin investigation, extended over a long succession of years and based on a greater material than any other investigation, has undoubtedly been of very great importance in this field to the generally accepted opinion about the superiority of breast feeding. A study planned and carried out like this one, nevertheless, must incur considerable sources of error. The material was collected by the physicians' reporting the kind of nourishment on the death certificates. In a great number of cases, however, such information was lacking. As an example it can be mentioned that in 1915, 4362 infants under one year of age died in Berlin. Of these 14.8% were entirely breast fed: 34.6% had received cow's milk exclusively; 18.2% had received mixed feeding; and in the remaining 32.4% there was no information about the kind of nourishment.

Table 1. Nourishment and infant mortality in Berlin (according to *Thiemich* and *Bessau*). Before reaching the next month of age the following number died out of each thousand infants.

	Breast fed infants			Bottle fed infants		
	1885–86	1895–96	1906	1885–86	1895–96	1906
1 month	22.4	19.6	22.4	142.0	111.9	59.1
2 months	9.0	7.3	7.9	82.7	58.7	31.3
3 months	6.8	4.3	4.3	72.2	49.7	27.3
4 months	6.4	3.6	2.4	61.8	46.6	22.1
5 months	5.3	2.6	1.7	57.1	37.0	18.5
6 months	4.9	2.5	2.2	50.7	31.0	16.1
7 months	4.7	2.5	1.4	46.5	27.7	14.1
8 months	4.5	2.3	1.8	40.8	24.1	12.2
9 months	5.3	2.0	2.1	33.3	21.3	10.2
10 months	5.4	3.8	1.5	29.5	19.1	9.2
11 months	6.3	3.1	1.3	24.9	16.7	8.0
12 months	—	3.6	1.5	—	14.6	8.0
Average	8.4	6.0	6.3	54.1	35.8	23.6

Fig. 1, taken from *Huber* [1919], shows that the mortality rate in the breast fed infants up to and including the eleventh month was only about 30% of the mortality in non-breast fed infants. The

figures originating from this material from Köln thus showed like the Berlin figures from the same time a three times greater risk of death for bottle fed than for breast fed infants.

Fig. 1. Mortality rate in Köln in 1919 (*Huber*) of breast fed infants (striped area) and bottle fed infants (line).

If it is right that the breast fed infants formerly had a much better prognosis it is, however, not sure that the situation is the same now. Artificial feeding has improved considerably. Investigations performed in recent years point to this fact. *Powers* and *Lawrence and assoc.* give a good survey of the progress in this field. *Nelson* points out "until recent years there was a much higher mortality rate among artificially fed infants than among breast-fed ones. Although the mortality rate of artificially fed infants has been greatly reduced, it is still higher and especially so among the lower socio-economic groups, than it is among breast-fed infants".

Table 2. The monthly mortality rate with regard to the feeding (according to *Woodbury*).

Month of life	Calculated mortality per month and per 1000 infants			
	All types of nourishment	Entirely breast feeding	Mixed	Entirely artificial feeding
1. First	44.8[1]	16.9	36.4	54.7
2. Second	9.3	5.8	14.7	24.6
3. Third	8.1	3.7	12.9	21.2
4. Fourth	8.0	3.4	9.0	19.2
5. Fifth	7.7	3.3	5.7	18.1
6. Sixth	7.4	2.1	5.9	17.7
7. Seventh	6.3	1.9	4.0	14.1
8. Eighth	5.8	2.9	3.3	11.3
9. Ninth	5.7	3.2	2.9	10.7
10. Tenth	5.3	3.8	2.3	9.3
11. Eleventh	3.9	2.4	2.5	6.0
12. Twelfth	4.5	4.4	2.7	6.4

[1] Death risk per 1000 infants is 21.5. 545 infants died without having received any food.

In *Woodbury*'s work [1926], based on more than 20,000 infants, the result was the same, with a lower mortality in breast fed than in bottle fed infants. The mortality rate is shown in Table 2.

Thus it turns out that death risk in the second month was 4.9 times greater for bottle fed than for breast fed infants. Corresponding figures were 5 for the 3rd -7th month, and 4 for the 8th month. *Woodbury* has also calculated the difference in mortality in such a way that he compared the real number of deaths among the bottle fed infants with the number of deaths that ought to have occurred if there had been the same mortality rate as among the breast fed ones. Instead of the 1047 deaths that really occurred, there should have been only 268.7 deaths. The real mortality rate was in other words almost 4 times greater than the calculated one.

A later investigation was published in 1934 by *Grulee, Sanford* and *Herron* and in 1935 by *Grulee, Sanford* and *Schwartz*. In the introduction to the former work the authors point out that artificial feeding in the United States has gotten such a fair start that "there has grown up the idea that artificial formulae can safely replace breast milk without any detrimental results to the child". The material consisted of 20,061 case records from the Infant Welfare Society of Chicago during 1924–29. The result showed a considerably higher mortality and morbidity among the bottle fed infants. 66.1% of the deaths occurred among the bottle fed infants, 27.2% among the partially breast fed ones, and only 6.7% among the breast fed infants. This investigation has been much criticized. Thus *Durand* (cit. *Grulee* [1934]) emphasized that the material consisted of children from the slum districts, which led to a selection from lower social groups, "but with intelligent mothers and direction from capable pediatricians as to the food, I do not think that the mortality today is higher on artificial food than it is on breast milk".

Some studies on the mortality and morbidity in breast fed and bottle fed infants in relation to different diseases have been published.

In ancient times a considerable number of the deaths during infancy were caused by acute and chronic nutritional disturbances. *Schlossman, Biedert, Heubner* and others have emphasized that these infants were almost exclusively bottle fed. The seasonal variation, with a concentration of cases during the summer (*Schlossman*), speaks for itself. This summer mortality has disappeared through improved infant care (*Meyer-Delius*).

Woodbury, using a statistical method, found that there was an 11 times greater mortality among bottle fed infants than among breast fed ones who died from gastrointestinal diseases up to the age of 9 months. All other causes of death showed the same tendency, though less pronounced. Thus, in the group of respiratory diseases the mortality rate in bottle fed infants was on the average 85% higher.

One way to examine the frequency distribution in breast fed and bottle fed infants is to study the distribution among infected children admitted to a hospital for infectious diseases. In 1942 such an examination was performed by *Ebbs* and *Mulligan* on 1500 infants admitted to the Hospital for Sick Children in Toronto. The number of breast fed infants among those with acute infectious diseases turned out to be less than half as great as at the child welfare centers of the city. Any significant difference in mortality among breast fed infants and bottle fed infants could, however, not be demonstrated. This difference in morbidity of breast fed infants and bottle fed infants, especially in respiratory diseases, has been stressed still more by *Cruickshank* [1945], and *Stevenson* [1947].

Except for the still birth and neonatal mortality (*Wallgren*, *Gyllenswärd*), which have not showed any decreasing tendency in the last decades, it is a fact that the mortality rate during the rest of infancy has been steadily diminishing. This is the case in all countries where the infant mortality has been the object for continuous statistic investigation.

1908 *Johannesen* and, 25 years later, *Schiötz* discussed the causes of this decreasing infant mortality. *Schiötz* says: "Diese Frage ist heute eigentlich als erledigt zu betrachten, denn die Tatsache der entschiedenen Überlegenheit der natürlichen und der Gefahren der künstlichen Ernährung wird nicht mehr in Frage gestellt".

It might not, however, be so simple as *Schiötz* suggested. It is well known that in America the infant mortality has decreased at the same time as the artificial feeding has become more common. A work from England by *Gordon* [1942], points in the same direction.

Woodbury emphasized that the mortality among bottle fed infants, dying from prematurity and malformations during the neonatal period, when different feeding could not have had any effect, was 10.7 times greater than among the breast fed infants. As is mentioned in Table 2, 545 infants that died without having received any food were excluded. The material thus consists of only 471

deaths, which gives a mortality of 21.5 per thousand. It does not seem reasonable to give figures for the death risks during the first month of life as most of the infants who died during this month were not included.

Woodbury discussed the influence of (1) prematurity, (2) nationality, and (3) socio-economics. He was the first one who discussed maternal weakness and disease as causes for the high mortality in bottle fed infants. *Woodbury*, however, stressed that both prematurity and higher maternal morbidity were of comparatively small importance. If this part of the difference in mortality was taken away from the total the remainder was so great that it was proved that "the type of feeding had a cumulative effect during early months of life". Nationality and race did not seem to have any great influence.

In 1944, *Deeny* and *Murdock* used the same method of study as the present author (see below). Their material consisted of 554 infants who died and a control group of 477 living infants. A comparative examination showed that the number of deaths among the breast fed infants was significantly lower than among the living ones. Income, type of housing or sequence of the child seemed to have no influence on the number of breast fed infants. The author mentioned in the discussion of the results "that this lowered incidence of breastfeeding should be regarded more as an effect than as a cause of unsuccessful reproduction". Like *Woodbury*, *Deeny* and *Murdock* suspected that heredity could be of importance.

Finally the *socio-economical conditions* have in different quarters been subjected to analyses.

Already in ancient times (*Kriege* and *Seutemann*, *Seutemann*) it was known that bottle fed infants among the economically poorest individuals had a higher mortality than those among the well situated. *Neumann* showed in an examination of dwelling standard and death risk by different feeding systems that the mortality both in breast fed infants and in bottle fed ones diminished with higher standard of life. The high mortality among the bottle fed infants also diminished with improved nursing. In 1912 *Seutemann* showed that the infant mortality in a group of families with an annual income of more than 1800 marks was for the bottle fed infants 11.2% and for the breast fed ones 5.8%. Corresponding figures in the group with an annual income less than 1800 marks were 27.4% and 9.6% respectively. *Woodbury* found, more than 20

years later, still some difference yet less pronounced. In the group "under $ 550" the death risk was 6.3 times greater for bottle fed than for breast fed infants. For the group "$ 1250 and over" the risk was only 4.1 times as high.

A study of the literature on the *frequency of naturally and artificially fed infants* is complicated because the criterion of breast fed and bottle fed infants is not uniform. In certain works, for example *Grulee and coworkers*, there is no account of what they mean with breast fed and bottle fed infants, and in other investigations, for example the Berlin investigation, the number of cases without information of the kind of feeding is considerable.

Table 3 is taken from *Prinzing* and shows the decreasing nursing tendency in Germany at the end of last century.

Table 3. Way of feeding in Berlin (according to *Prinzing*).

	1885	1890	1895	1900	1905	1910
Mother's milk	55.2	50.7	43.1	31.4	31.2	30.5
Nurse milk	2.7	2.2	1.4	0.7	0.6	0.4
Mother's milk + addition	6.7	4.8	9.9	14.4	4.2	3.7
Artificial feeding only	33.9	42.3	45.3	49.7	63.7	62.6
Unknown	1.5	0.0	0.2	3.8	0.3	2.8

Similar results are reported from England regarding the last 20 years (*Mackintosh*).

Woodbury [1926], reported that during 1911–1915 78.3% were breast fed up to 3 months, 5.9% got mixed feeding, and 15.8% got cow's milk only. At 6 months of age corresponding figures were 51.7%, 29%, and 19.3%. The figures for the 11th and 12th months of living were surprising as, contrary to Swedish circumstances, not less than 13.3% of the one year old infants were entirely breast fed.

From the United States, furthermore, the figures by *Grulee and coworkers*, from 1934, are available showing that 48.5% of the infants were breast fed, 8% bottle fed and the remainder, or 43%, got mixed feeding.

Mackintosh [1944], reported that about 50% of the infants were breast fed up to 3 months of age and 30–40% up to 6 months of age. Three years later *Williams* [1947] published figures from Oxford showing that not less than 73% were breast fed until 3 months of age, 58% until 5 months of age, and 47% until 7 months of age.

From Denmark *Herz* [1930], reported that only 45% of the infants were breast fed up to 2 months of age and 15% up to 6 months of age.

From Sweden recent figures are available, built on a great material from all child welfare centers of the country. The number of infants controlled by child welfare centers has the last years been remarkably great in our country. *Lichtenstein* reported that in 1945 85%, or 111,000 out of 134,000 infants were regularly controlled by the child welfare centers. *Ström* reported, in 1946, that in Sweden in 1938-1944 70-75% of the infants were breast fed up to 2 months and 42–46% up to 6 months. In 1944 corresponding figures for Stockholm were 82% and 56% respectively. As these figures concerned only infants continuously controlled at the child welfare centers, the rest of the infants were not included. These represent a poorer social selection. *Rudberg* showed on a material from Stockholm 1936–37 including an inquiring on the method of feeding among 2013 mothers that infants controlled by a child welfare center showed a higher percentage of nursing than an uncontrolled material. *Rudberg* also demonstrated that the length of the nursing time during the last 15 years had not decreased.

As the Scandinavian countries are characterized by a low infant mortality and relatively greater number of breast fed infants, and as no reports on infant mortality in Scandinavia in recent years are available this work has been planned.

Methods

The method used in this investigation is different from those mostly used in preceding works. Earlier a follow up examination from birth until one year of age was made also including the number of deaths. In order to avoid sources of error always connected with such a method a "sampling method" has been used. Healthy individuals chosen at random were collected. Figures from such a study are representative, because they have been applied upon all newborns. In order to get information about the deaths, all infants who died from the 2nd up to the 12th month were collected. The mortality in the first month of life is of less interest as it could not be caused by the feeding. The method used in this study is a combination of "sampling" and complete statistics.

Deeny and *Murdock* used the same method in 1944, of which,

however, was not known by the present author when this study was begun.

Material

The material in this study consists of all infants who died in Stockholm during the years 1943–47 at the above mentioned age, i.e. from the second to the twelfth month inclusive. This material (see Table 6a) was received from the Medical Officer of Stockholm. Reports about the infant feeding were then obtained from case records and by home visits. Mostly very conscientious information was given in the records. The mothers were questioned about those for whom information of this kind was lacking, and about those who had died at home. The size of the material can be seen in Table 6a. Reports could not be obtained for a small group of 18 infants born during the years 1943–1947, who had moved away from Stockholm.

The control material was collected in the following way. In the Royal Central Office of Statistics a so-called birth book contains the names of all living infants in Stockholm. Taking infants from this book at random a control material representative of Stockholm was collected, consisting of 317 infants born in Stockholm during 1943–1947. Home visits were then made in the same way as for the deaths.

To get information about the family income extracts from the assessment books were procured.

Results

Frequency of different sorts of feeding. A division into entirely artificially fed infants covers only a part of the material. The distribution of different feedings among the living infants month by month is shown in Fig. 2a and Table 5a. One finds, for example, that only 28% were entirely breast fed while 20% were exclusively artificially fed at the age of 6 months. Not less than 52% got mixed feedings of varying kinds. If one wishes to work up the material statistically, it is thus necessary to present certain criteria for what is meant by "breast fed" and "bottle fed" infants, respectively.

In the present study the following classification was used:

1. Those who were entirely breast fed or had got an insignificant addition of, at most, ¼ cow's milk mixture up to and including the

sixth month were counted as breast fed infants. The remainder were regarded as non-breast fed infants.

These last were in turn divided into the following groups:
2. Those who had received cow's milk exclusively during the whole time or, in any case, from the end of the first month were regarded as bottle fed infants. Also in this group were those who had got only an insignificant addition of cow's milk during the first quarter and then at least half cow's milk.
3. The remainder were classified as infants given mixed feedings.

The bottle fed infants were divided into two groups. In the first were those who received cow's milk exclusively from the second month onwards, i.e., breast fed for at most one month. This group was called "bottle fed infants of first degree", while the others were called "bottle fed infants of second degree". This latter division was only used in calculating the mortality.

It now appeared that in the control group of living infants, 59.3% were breast fed and 15.8%, bottle fed. The remainder got mixed feedings. It should be mentioned that the occurrence of bottle fed infants of first degree according to the definition given above was only 5.05%. The figures are regarded as very good when compared with those of other countries from which information is available.

A comparison with figures for Stockholm in 1944 (*Ström*) (82% breast fed infants up to 2 months of age and 56% up to 6 months of age) gave the following results. Table 5a and Fig. 2a show that 83.6% of the control material consisted of infants who were breast fed up to 2 months of age, a figure which corresponded well with that of *Ström*. On the other hand it appeared that only 28.4% were *entirely* breast fed up to 6 months of age. If one, however, regarded all 6-month-old infants who received any breast milk at all as breast fed infants (which could be well justified), quite another figure of 80.4% was obtained. The figure given by *Ström*, 56%, lies half-way between these extremes. The result shows how necessary it is to define clearly the concepts used before the figures about feeding are given. The figures given by *Ström* were not comparable because his entire material was controlled at child welfare centers.

Infants with and without supervision. In this material infants who visited child welfare centers were compared with those who did not (Table 4 a and b).

Table 4a. Percentage distribution according to visits at the child welfare centers, divided into feeding groups.

Visited child welfare centers	Breast fed infants		Mixed fed infants		Bottle fed infants		Total	
	Number	%	Number	%	Number	%	Number	%
Regularly	121	64.4	48	60.7	24	48.0	193	60.9
Irregularly or only once .	6	3.2	7	8.9	5	10.0	18	5.7
Never (the remainder) . .	61	32.4	24	30.4	21	42.0	106	33.4
Total	188	100.0	79	100.0	50	100.0	317	100.0

It appeared that 64.4% of the breast fed infants but only 48% of the bottle fed infants regularly visited child welfare centers. The breast fed infants were, in other words, greater in number among those who visited child welfare centers and less in number among the non-supervised. Among the later 35.6% were breast fed while 52% were bottle fed. On an average 80% of the infants visited child welfare centers regularly in Stockholm during these years. The figures show that the child welfare centers attracted mothers who took rather good care of their infants, which was to be expected. Those who visited the child welfare centers were naturally influenced to nurse their infants. Both these facts surely played a rôle in the sort of feeding the infants received.

Table 4b shows the number of "dead infants" previously supervised or non-supervised. The difference was still more stressed in this material, indicating that the dead infants tended to be less well cared for. Among the breast fed infants who died during the 2nd–12th month 30.2% had regularly visited child welfare centers, while among the bottle fed infants, the corresponding figure was 21.3%. For "the dead" during the 3rd–12th month the corresponding figures are 42.3% and 21.7%, respectively.

It is thus evident that one gets a far too optimistic picture when the investigation is limited to material from child welfare centers.

It appears from the above that the method worked out for studying feeding in our country should give information about the conditions during a fixed time in a representative cross section of the population. All groups of the population, independent of social classes or manner of care, occur in a normal sort of distribution.

Table 4b. Percentage distribution of the dead infants with regard to visits at child welfare centers and divided into feeding groups. (For number of cases, see Table 6a.)

Visited child welfare centers	Breast fed infants	Mixed fed infants	Bottle fed infants	Total
	Dead in 2nd–12th month			
Regularly	30.2	27.8	21.3	26.9
Irregularly or only once .	12.1	19.4	9.3	12.3
Not at all (the remainder)	57.7	52.8	69.3	60.8
Total	100.0	100.0	100.0	100.0
	Dead in 3rd–12th month			
Regularly	42.3	26.5	21.7	32.0
Irregularly or only once . .	14.1	20.6	6.7	12.8
Not at all (the remainder)	43.6	52.9	71.7	55.2
Total	100.0	100.0	100.0	100.0

Mortality among breast fed and bottle fed infants. As an introduction a survey of the feeding during different months was made. Table 5a and b gives figures for breast fed infants and bottle fed ones of different degrees during the different months. Table 5a gives figures for the infants who survived the first year. Table 5b gives figures for the infants who died between the end of the first month and the twelfth month. It appears that the occurrence of breast fed infants was higher among the living than among the dead infants. As an example, in the 4th month, 65% were entirely breast fed among the living infants, while among the dead infants the corresponding figure was 44%. On the other hand, 10% of the living infants in the same month were entirely bottle fed, while the corresponding figure was 21% for the deaths. These figures give a strong impression that bottle fed infants are more likely to die than breast fed ones. Figs. 2a and b show this fact graphically. These curves and tables give, however, no conception of how much greater is the risk for bottle fed infants.

First the figures for the living infants were used to distribute all infants born in Stockholm during these years in the same proportions as in the control material. The deaths were then placed in relation to these groups. Table 6a shows the death risks for different feedings. A comparison between breast fed and non-breast fed infants showed a difference in mortality for the 2nd–12th month. This difference was accentuated when the comparison was limited to the 3rd–12th month.

Table 5a. Distribution of the control material with regard to kind of feeding in each month of the first year of life (in per cent of all cases).

Months of the first year of life	Percentage distribution with regard to kind of feeding						
	Entirely breast fed	Insignificant addition of cow's milk	¼ addition	¼–½ addition	½–¾ addition	Entirely artificial feeding	Total
1	89.9	0.6	3.5	1.6	1.3	3.2	100
2	83.6	1.9	4.7	3.8	0.9	5.0	100
3	77.0	1.9	6.9	5.4	1.9	6.9	100
4	64.7	7.3	8.8	4.7	4.1	10.4	100
5	46.4	11.7	14.8	7.3	6.3	13.6	100
6	28.4	9.5	21.8	12.0	8.8	19.6	100
7	9.5	9.1	19.2	19.2	13.6	29.3	100
8	3.8	4.4	9.5	18.0	21.5	42.9	100
9	1.6	1.9	3.2	9.1	18.9	65.3	100
10	0.3	0.6	2.5	3.5	8.5	84.5	100
11	0.3	—	0.6	2.8	2.5	93.7	100
12	0.3	—	0.3	0.3	2.2	96.8	100

Table 5b. Infants dead during the 2nd–12th month, distributed with regard to the kind of feeding in each month (in per cent of number of living infants in each month of age).

Months of the first year of life	Percentage distribution with regard to kind of feeding						
	Entirely breast fed	Insignificant addition of cow's milk	¼ addition	¼–½ addition	½–¾ addition	Entirely artificial feeding	Total
1	80.2	1.3	4.4	2.6	1.8	9.7	100
2	62.3	5.0	4.0	10.6	1.5	16.6	100
3	51.7	3.3	9.3	13.2	4.6	17.9	100
4	44.3	3.3	6.6	19.7	4.9	21.3	100
5	36.5	8.3	13.5	7.3	10.4	24.0	100
6	20.8	10.4	19.5	13.0	6.5	29.9	100
7	5.5	9.1	16.4	18.2	14.5	36.4	100
8	—	7.5	5.0	20.0	17.5	50.0	100
9	—	—	6.3	9.4	25.0	59.4	100
10	—	—	—	4.2	16.7	79.2	100
11	—	—	—	—	7.7	92.3	100
12	—	—	—	—	—	100.0	100

With regard to the mean error the first mentioned difference is only probable, while that for the 3rd–12th month is statistically significant. Proceeding with the comparison it appears that the

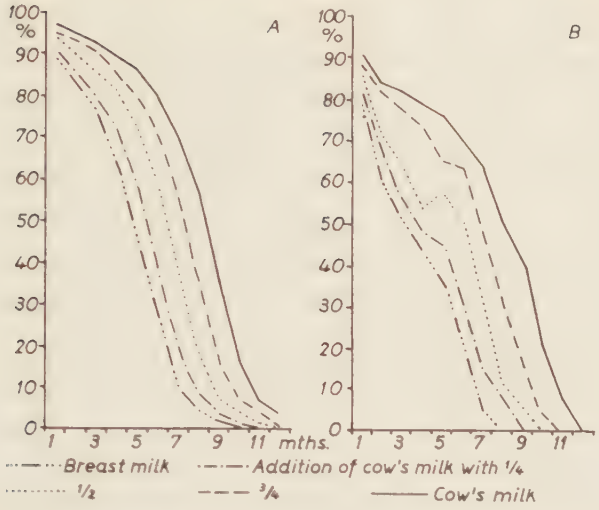


Fig. 2a and *b*. Feeding during the first year of life. a) Control material. b) Infants who died during the 2nd–12th month.

Table 6a. Death risks for infants during the 2nd–12th and 3rd–12th month (per thousand) for different kinds of feeding

Kind of feeding	Death risks during the 2nd–12th month			Death risks during the 3rd–12th month		
	Number of born reduced by number of deaths during the first month	Number of deaths	Deaths per thousand	Number of born reduced by number of deaths during the first month	Number of deaths	Deaths per thousand
Breast fed	40,173	186	4.63 ± 0.34	40,115	128	3.19 ± 0.28
Non-breast fed	27,565	169	6.13 ± 0.47	27,539	143	5.19 ± 0.43
Difference between non-breast fed and breast fed	—	—	1.50 ± 0.58	—	—	2.00 ± 0.51
Non-breast fed:						
Mixed feeding	16,881	52	3.08 ± 0.43	16,878	49	2.90 ± 0.41
Bottle fed	10,684	117	10.95 ± 1.01	10,661	94	8.82 ± 0.90
Out of which:						
a) exclusively cow's milk from the 2nd month onwards	3,419	58	16.96 ± 2.21	3,408	47	13.79 ± 2.00
b) remaining bottle fed infants	7,265	59	8.12 ± 1.05	7,253	47	6.48 ± 0.94
Difference between a and b	—	—	8.84 ± 2.45	—	—	7.31 ± 2.21

bottle fed infants had the greatest mortality, namely 11 per thousand. When classifying only the most extreme cases ("bottle fed infants of first degree") as bottle fed infants, according to the definition given above the mortality was, as shown in the table, 17 per thousand for the 2nd–12th months and 14 per thousand for the 3rd–12th months. Quite surprisingly the breast fed infants come next with 4.6 per thousand. Finally, those with mixed feeding had a mortality of only 3 per thousand. That the mortality is so low naturally depends upon the fact that those who died during the first month have been excluded. It is well known that the greatest mortality (the neonatal mortality) occurs during this month. It is caused by malformations, prematurity and general weakness. The most important fact is that this mortality cannot depend on the sort of feeding. Because of that the deaths were also excluded during the second month. In Table 6a the death risks for different feeding groups during the 3rd–12th month are shown. Here, the breast fed infants also show a higher figure than the mixed fed. but the difference is so small (0.29 ± 0.40) that, with regard to the mean error, it may be caused by chance. The bottle fed infants show the highest figure, 9 per thousand. According to these figures the bottle fed infants had a mortality three times greater than the breast fed ones. The figure corresponds rather well with those in reports by both *Huber* and *Woodbury*. When the second month was included, the difference was not quite so great. In that case the mortality of the bottle fed infants was a little more than twice that of the breast fed ones.

An analysis of the death risk for different feedings quarter by quarter is given in Table 6b. The first quarter is represented by the second and third months, the first month being excluded in this study. The following results were found. Infants with mixed feeding had the lowest mortality during the second and third months. During the following three months it was almost the same as for the breast fed infants, and during both the last quarters it was somewhat higher than for the breast fed infants. Regarding the bottle fed infants, their high mortality was marked only during the first three quarters. The difference in mortality between breast fed and mixed fed infants during the second month was statistically significant (difference $= 1.26 \pm 0.21$ per thousand), as was also the difference between bottle fed and breast fed infants during the same month (difference $= 0.71 \pm 0.23$ per thousand).

Table 6b. Death risks during different months of age among infants with different feedings.

Kind of feeding	Deaths, age in months	Number of births, reduced by number of deaths during the first month, and during each month thereafter	Number of deaths in each month of age	Deaths per thousand	Per thousand each month
Breast fed	2nd	40,173	58	1.44	1.44
	3rd	40,115	27	0.67	0.67
	4th–6th	40,088	45	1.12	0.37
	7th–9th	40,043	30	0.75	0.25
	10th–12th	40,013	26	0.65	0.22
Mixed feeding	2nd	16,881	3	0.18	0.18
	3rd	16,878	4	0.24	0.24
	4th–6th	16,874	16	0.95	0.32
	7th–9th	16,858	16	0.95	0.32
	10th–12th	16,842	13	0.77	0.26
Bottle fed	2nd	10,684	23	2.15	2.15
	3rd	10,661	24	2.25	2.25
	4th–6th	10,637	42	3.95	1.32
	7th–9th	10,595	19	1.79	0.60
	10th–12th	10,576	9	0.85	0.28

Fig. 3 presents graphically the death risks for different feeding systems. The fact mentioned above, that *the markedly high mortality of the bottle fed infants during the first three quarters tended to disappear later, is obvious here.*

Fig. 3. Graphic presentation of the death risk for different feeding systems.

Of most importance in the results obtained is the fact that a markedly high mortality continued among the bottle fed infants in spite of the general decrease in infant mortality rates, and in spite

of progress in modern nutritional physiology which has led to considerable improvement in artificial feeding. The death risk for a bottle fed infant in Stockholm is still from two to three times that for a breast fed infant. The infants with mixed feeding in our material seemed to be very favorably situated. This result, which is rather surprising, could be explained by the breast fed infants being especially weak children. Other factors of selection could also play a rôle. For example, the mothers of the infants with mixed feeding may have been especially energetic in trying to continue breast feeding and thus may have given their infants excellent care.

Different causes of death. The causes of death were divided into the following five groups:

1. Congenital diseases.
2. Infectious diseases.
3. Respiratory diseases.
4. Intestinal diseases.
5. Other diseases.

Table 7 shows the percentage distribution in the different groups of the causes of death and also the death risk in the different causes of death with regard to the three kinds of feeding. The division of the material into so many small classes may have sometimes led to results that were not meaningful statistically. It appears from the table that among infants dying during the 2nd–12th months intestinal diseases seemed to be most common proportionately in bottle fed infants. No significant difference seemed to exist between breast fed and mixed fed infants. The risk figures gave a better idea of the situation. It then appeared that the risk of death from congenital diseases seemed to be about twice as great among bottle fed infants as among breast fed ones, but it was still lower among infants with mixed feeding. Regarding infectious diseases the difference was not so great, but it pointed in the same direction. With regard to respiratory diseases, a considerably higher risk was found for the bottle fed infants. An at least equally marked difference existed in intestinal diseases and in other diseases. If the comparison was limited to infants dying during the 3rd–12th month there was of course a somewhat lower risk of death from congenital diseases, but the risk in the other groups of diseases also decreased. The results showed that the death risk was greater among bottle fed than among breast fed infants in all groups of diseases studied.

Table 7. Percentage distribution with regard to causes of death and the death risk (per thousand) in the different groups of diseases among infants with different feedings. For deaths during the 2nd–12th month and for deaths during the 3rd–12th month.

Infants who died during	Causes of death	Breast fed			Mixed feeding			Bottle fed		
		Deaths %	Number	Death risk per thousand	Deaths %	Number	Death risk per thousand	Deaths %	Number	Death risk per thousand
2nd–12th month	1. Congenital diseases	27.6	32	1.28	22.2	8	0.68	20.3	15	2.22
	2. Infectious diseases	25.0	29	1.16	19.4	7	0.60	13.5	10	1.48
	3. Respiratory diseases	36.2	42	1.68	50.0	18	1.54	39.2	29	4.29
	4. Intestinal diseases	5.2	6	0.24	8.3	3	0.26	14.9	11	1.63
	5. Other diseases	6.0	7	0.28	—	—	—	12.2	9	1.34
	Total	100.0		4.63	100.0		3.08	100.0		10.95
3rd–12th month	1. Congenital diseases	25.6	20	0.82	20.6	7	0.60	20.3	12	1.79
	2. Infectious diseases	24.4	19	0.78	17.6	6	0.51	11.9	7	1.05
	3. Respiratory diseases	39.7	31	1.27	52.9	18	1.54	37.3	22	3.29
	4. Intestinal diseases	2.6	2	0.08	8.8	3	0.26	17.0	10	1.50
	5. Other diseases	7.7	6	0.25	—	—	—	13.6	8	1.20
	Total	100.0		3.19	100.0		2.90	100.0		8.82

The age of the mother. From Table 8 it appeared that the age of the mother was rather low among the dead infants compared with the living ones. As an example, it can be mentioned that 14.2% of the mothers of the living infants were 20–24 years old, while corresponding figure for the mothers of the dead infants was 25.5%. This is an expression of the well known fact that infants of younger mothers show a somewhat greater death risk. Of especial interest in this connection, however, is the influence of different feedings. Table 9 presents the mean and the median age for mothers in the different feeding groups. No significant difference could be found. The possibility that bottle fed infants were recruited from younger mothers, which should have explained a part of the difference between the groups, is thus excluded.

Table 8. Percentage distribution of the age of the mother in the control material and among infants who died during the 2nd–12th month.

The age of the mother at delivery	Percentage distribution in	
	Control material	Deaths during the 2nd–12th month
15–19	0.6	2.0
20–24	14.2	25.5
25–29	32.3	28.4
30–34	30.7	22.1
35–39	17.4	17.2
40–44	4.7	4.4
45 and more	—	0.5
Total	100.0	100.0

Table 9. The age of the mother at delivery, among infants who died during the 2nd–12th month, distributed with regard to feeding.

Feeding	The age of the mother at delivery	
	Average	Median
Breast fed	29.2	28.4
Mixed feeding	29.5	26.8
Bottle fed	29.3	28.5

Birth weight and feeding. It is well known that infants with low birth weights have a greater death risk, which partly depends on their short time in utero. It is shown in Table 10 that those who weighed 2500 grams or less had a much greater death risk than the

other weight classes. The death risk seemed to decrease continuously with increasing birth weight, at least up to the highest weight groups. The difference in death risk for the three weight groups was, however, insignificant. This difference was not related to the feeding. Among infants with a birth weight up to 2500 grams, a high number were breast fed, namely 60%, while the breast fed infants in the other weight groups numbered about 50%. Inversely, the occurrence of bottle fed infants was especially low in the lowest weight class, namely 23.5%, while in the other weight classes this figure was about 30–35%.

The high mortality for breast fed infants among the prematures depended undoubtedly, on the fact that the group included the lowest weight group (less than 1500 grams).

Difference in infant mortality between different social groups. It has been known for a long time that infant mortality is very different in different social groups. Infant mortality has even been used as a measure of social differences. The information used was received in the form of assessed income. Although this does not always correspond to the real income it might be quite satisfactory. Information was also received about the number of rooms in which each family lived. The number of rooms might not run quite parallel with the social situation but one ought to get an idea of family circumstances with the aid of reports about dwelling size.

First, the size of the death risk during the 2nd–12th month in different income groups (see Table 11) was investigated. Reports of the incomes of the families of 1.6% of the living infants and of 18% of those who died could not be obtained. This already indicates that it was especially those children in families with low incomes who died. Table 11 shows the results. Among those where the family's annual income was 5000 Swedish crowns or less, the death risk was 9 per thousand, while the mortality rate was 4.4 per thousand among those where the family's annual income was 5,000–10,000 Swedish crowns. Those whose families were in the highest income group, i.e. 10,000 crowns or more, had a death risk of 3 per thousand. The result corresponds well with previous statements. One may now imagine that this difference depends on the fact that breast fed infants were less common among those with low incomes. This, however, was not the case. As is evident from the table, there was no significant difference in the distribution of feeding groups among the different income classes.

Table 10. Death risk (per thousand) among infants with different kinds of feeding and different birth weight, and percentage distribution among feeding groups with regard to birth weight.

Birth weight, grams	Living	Kind of feeding	Deaths	Infants who died during the 2nd–12th month	
				Death risk per thousand	Distribution among feeding groups, %
2,500 or less	855	Breast fed	41	47.95	60.3
	427	Mixed feeding	11	25.76	16.2
	641	Bottle fed	16	24.96	23.5
Total	1,923		68	35.36	100.0
2,510–3,000	3,633	Breast fed	29	7.98	47.5
	1,068	Mixed feeding	10	9.36	16.4
	1,068	Bottle fed	22	20.60	36.1
Total	5,769		61	10.57	100.0
3,010–3,500	14,317	Breast fed	60	4.19	53.1
	7,052	Mixed feeding	16	2.27	14.2
	3,419	Bottle fed	37	10.82	32.7
Total	24,788		113	4.56	100.0
3,510–4,000	13,249	Breast fed	40	3.02	51.9
	5,983	Mixed feeding	11	1.84	14.3
	4,060	Bottle fed	26	6.40	33.8
Total	23,292		77	3.31	100.0
4,010 or more	8,120	Breast fed	20	2.46	55.6
	2,350	Mixed feeding	5	2.13	13.9
	1,496	Bottle fed	11	7.35	30.6
Total	11,966		36	3.01	100.0

A comparison between these figures and those given by *Rietz* [1930] regarding the years 1915–1922 (Table 11 and Fig. 4) is of interest. In the income class of less than 4000 crowns the mortality in the 2–12 month age group was 24.8 per thousand. In the class getting 6,000–10,000 crowns, the mortality was 12.2 per thousand and in the class getting more than 10,000 crowns it was 2.9 per thousand. In the present material, corresponding figures are 8.88, 4.37 and 2.99 per thousand respectively. Disregarding the monetary fall, the infant mortality in Sweden seems to have kept noticeably

constant in the highest social group throughout the last 30 years. If one takes into consideration the monetary deterioration that has taken place in reality, there is reason to suppose the infant mortality in the corresponding social group may be less than 2.9 per thousand, and that there only appears to be the agreement mentioned above. However it may be evident that infant care in the prosperous part of the population 30 years ago was already of a good and high standard. It is true that there is still a significant difference between different social groups, but considerable improvement has undoubtedly taken place. In the years 1915–1922 the death risk for the infant was almost 8 times greater for the poor than for the well-to-do. Nowadays, the corresponding figure is barely three. The result is to a certain degree an illustration of the importance of prophylactic infant care and a testimony that this has an effect among the lower social groups especially.

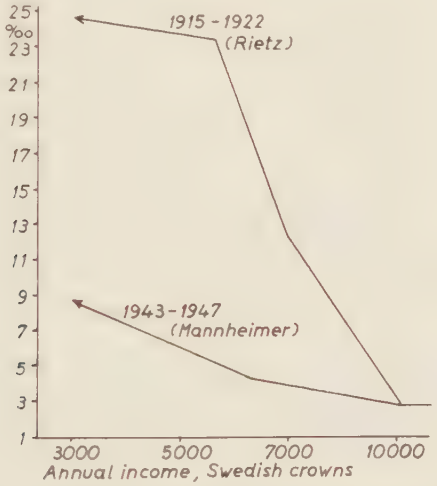


Fig. 4. Infant mortality in different income classes. A comparison.

Looking at the percentage distribution, there were no obvious differences between the causes of death in the different income groups (see Table 12). A truer picture is obtained by studying the death risks for different causes of death.

The death risk in congenital diseases was much greater for those in the low income groups than for those in the moderate or high income groups. The same was also the case regarding the other causes of death. The death risk from intestinal and other diseases

was, however, only significantly higher for those in the low income groups, while the risk from infectious and respiratory diseases showed a marked difference.

Table 11. Percentage distribution according to kind of feeding among the control material and among infants who died during the 2nd–12th months with regard to their parents' annual income, and death risks (per thousand) in corresponding groups.

The parents' annual income in Sw. cr.	Kind of feeding	Percentage distribution according to feeding among		Death risks per thousand
		the control material	those who died during the 2nd – 12th month	
5,000 or less	Breast fed	63.0	58.2	8.31 ± 0.86
	Mixed feeding	19.8	9.5	4.32 ± 1.11
	Bottle fed	17.3	32.3	16.78 ± 2.33
	Total	100.0	100.0	8.98 ± 0.71
5,000–10,000	Breast fed	57.3	55.0	4.20 ± 0.46
	Mixed feeding	28.7	15.5	2.35 ± 0.49
	Bottle fed	14.0	29.5	9.21 ± 1.38
	Total	100.0	100.0	4.37 ± 0.36
10,000 or more	Breast fed	60.8	39.6	1.94 ± 0.45
	Mixed feeding	23.0	20.8	2.71 ± 0.86
	Bottle fed	16.2	39.6	7.29 ± 1.67
	Total	100.0	100.0	2.99 ± 0.43

Table 12. Percentage distribution according to causes of death and the death risk (per thousand) in the different groups of diseases among infants who died during the 2nd–12th month, divided with regard to their parents' annual income.

Causes of death	The parents' annual income, Swedish crowns					
	5,000 or less		5,000–10,000		10,000 or more	
	Deaths %	Death risk per thousand	Deaths %	Death risk per thousand	Deaths %	Death risk per thousand
1. Congenital diseases	27.7	2.49	22.1	0.97	28.0	0.84
2. Infectious diseases	26.5	2.38	11.7	0.51	16.0	0.48
3. Respiratory diseases	36.1	3.24	48.1	2.10	36.0	1.08
4. Intestinal diseases	4.8	0.43	7.8	0.34	12.0	0.36
5. Other diseases	4.8	0.43	10.4	0.45	8.0	0.24
Total	100.0	8.98	100.0	4.37	100.0	2.99

Finally, a comparison was made with regard to the dwelling size (see Table 13). The dead infants, proportionately, had a small dwelling more often. Of the living infants, 33.7% were in flats consisting of one room and kitchen or one room without kitchen, but 46.9% of the dead infants had lived in such dwellings. On the other hand, 5.6% of the living infants lived in five rooms and kitchen or more, but only 3.9% of the dead ones. These figures do not shed any more light on the problem; they only support what has already been found by calculating the risk figures for different income groups.

Weight and increase in weight during the first year of life. Although increase in weight does not belong to the problems in this study, the following reports were worked up. Figures for normal weight are given by *von Sydow* and *Broman*, *Dahlberg* and *Lichtenstein*. In analyzing the weight the latter standards were used. Mortality compared with birth weight was previously studied. To get an idea of the increase in weight up to the end of the 12th month only those infants could be used who had lived up to that age, and, furthermore, only those with a birth weight between 2500 and 4000 grams, in order to make possible a comparison with the normal figures mentioned above.

Table 13. Percentage distribution according to the size of dwelling among the control material and among infants who died during the 2nd–12th month.

Size of dwelling	Percentage distribution among	
	the control material	those who died during the 2nd–12th month
1 room without kitchen	2.2	13.4
1 room with kitchen	31.5	33.5
2 rooms without kitchen	0.3	1.7
2 rooms with kitchen	38.2	32.4
3 rooms without kitchen	—	0.6
3 rooms with kitchen	15.5	10.6
4 rooms with kitchen	6.6	3.9
5 rooms with kitchen	2.5	2.2
more than 5 rooms with kitchen	3.2	1.7
Total	100.0	100.0

The infants were divided with regard to the kind of feeding. It may be seen from Table 14 that a good agreement was found with the normal figures for the different feeding groups. The increase in weight was probably somewhat better for those with mixed

feeding than for the others. The difference was, however, so insignificant that it might have been caused by chance. The figures agreed very well with the normal figures given by *von Sydow* (in the whole material the difference was only 36 grams). They showed further that there was no real difference in the increase in weight with regard to the kind of feeding.

Table 14. Average weight at one year of age for the different feeding groups in the control material, given with the weighted average figures according to sexual distribution in corresponding birth weight and feeding groups.

Feeding	Birth weight grams	Average weight at one year of age for the control material	Weighted normal figures for the weight at one year of age	Difference normal figures — control material
Breast fed	2510–3000	9,484	9,455	— 29
	3010–3500	9,740	9,810	70
	3510–4000	10,153	10,195	42
Mixed feeding	2510–3000	9,655	9,375	— 280
	3010–3500	9,936	9,880	— 56
	3510–4000	10,091	10,090	— 1
Bottle fed	2510–3000	8,937	9,200	263
	3010–3500	9,725	9,808	83
	3510–4000	9,909	10,043	134
Feeding and birth weight groups, total	2510–4000	9,881	9,917	36

Furthermore, differences between the different income groups were looked for here. As is shown in Table 15, there were none. Taking the different birth weights into consideration the increase in weight was equal and was only 35–50 grams less than the normal figures.

It is thus obvious that neither income group nor kind of feeding seemed to influence the weight increases.

Table 15. Average weight at one year of age of infants in the control material with a birth weight between 2500 and 4000 grams (= the three middle groups taken together), distributed with regard to the parents' annual income; and normal figures, given with weighted average figures according to the sexual distribution and birth weight of the control material.

Annual income Sw. cr.	Average weight at one year of age for the control material	Weighted normal figures for the weight at one year of age	Difference normal figures—control material
5,000 or less	9826	9861	35
5,000–10,000	9909	9941	32
10,000 or more	9868	9917	49

Summary

The present study deals with the problem of mortality in a representative material of breast fed and bottle fed infants in Stockholm in 1943–1947.

1. Mothers who nursed their infants visited the child welfare centers more often than did other mothers. 64% of the breast fed infants visited child welfare centers regularly. The corresponding figure for the bottle fed infants was 48%. It is thus impossible to get a representative material by taking children from welfare centers only.

2. It appeared that the bottle fed infants still had the greatest mortality. The death risk for an artificially fed infant was 2 to 3 times that for a breast fed one. This high mortality rate for bottle fed infants was greatest in the beginning of infancy, decreasing greatly during the fourth quarter. The mixed fed infants in the 2nd and 3rd months of life seemed to be especially favorably situated.

3. The high mortality for bottle fed infants was found in all causes of death.

4. The age of the mother and the birth weight of the infant did not seem to influence the mortality in different kinds of feeding, except for the prematures, including a group of especially small and weak breast fed infants with high mortality.

5. The difference in infant mortality between different social groups was estimated partly by comparing infants from different income groups, partly by dividing the material with regard to the dwelling situation. There was a higher infant mortality in families with lower annual income and with smaller dwellings. However, formerly the difference was much greater than at present, and in the highest income group (an annual income of more than 10,000 Swedish crowns) the infant mortality was the same as 30 years ago. There is no significant difference in the distribution of feeding groups among the different income groups.

6. Finally, it appeared that neither different annual income nor different feeding influences the increase in weight of the infant to any great extent. The increase in weight in the present study shows good agreement with figures given previously by *von Sydow*.

Résumé

La présente étude est consacrée au problème de la mortalité d'un groupe représentatif de nourrissons nourris respectivement au sein ou au biberon, entre 1943 et 1947 à Stockholm.

1. Les mères allaitant leurs enfants consultaient les centres de puériculture plus souvent que les autres. 64% des nourrissons au sein y furent contrôlés régulièrement, tandis que le chiffre correspondant des nourrissons au biberon ne s'éleva qu'à 48%. De ce fait, il est impossible de se procurer un matériel statistiquement représentatif en ne considérant que les enfants des centres de puériculture.

2. Les nourrissons alimentés au biberon présentent toujours la mortalité la plus élevée. Le risque de mort pour un enfant nourri artificiellement est deux à trois fois plus élevé que pour un enfant nourri au sein. Ce taux de mortalité élevé des nourrissons alimentés artificiellement atteint son point culminant au début de la vie, et diminue rapidement pendant le quatrième trimestre. Les enfants soumis à une alimentation mixte semblent jouir d'une situation particulièrement favorable pendant les deuxième et troisième mois de la vie.

3. La mortalité élevée des nourrissons alimentés au biberon se retrouva dans toutes les causes de décès.

4. L'âge de la mère, pas plus que le poids de l'enfant à la naissance, ne semblèrent influencer la mortalité des divers groupes considérés, si ce n'est celle des prématurés, comprenant un groupe de nourrissons au sein particulièrement petits et faibles, à mortalité élevée.

5. La différence de mortalité infantile selon les catégories sociales fut estimée d'une part en comparant les nourrissons selon les classes de revenus de leurs parents, d'autre part en subdivisant le matériel en groupes correspondant aux conditions d'habitation. La mortalité infantile s'avéra plus élevée dans les familles dont les revenus annuels étaient bas, et dans celles dont les locaux d'habitation étaient petits. Cependant, cette différence s'était montrée autrefois beaucoup plus marquée qu'aujourd'hui et, parmi les familles aux revenus les plus élevés (revenu annuel de plus de 10 000 couronnes suédoises, la mortalité infantile se trouva la même que 30 ans plus tôt. La répartition selon le mode d'alimentation ne

montre pas de différence significative à l'intérieur des différentes classes de revenus.

6. Enfin, il apparut que ni les différences de revenus annuels ni celles du mode d'alimentation n'exercent d'influence notable sur l'augmentation du poids de l'enfant; cette dernière, dans la présente étude concorde d'une façon satisfaisante avec les chiffres donnés récemment par *von Sydow* [1940].

Zusammenfassung

Die vorliegende Untersuchung beleuchtet die Frage der Sterblichkeit der Brustkinder und Flaschenkinder in Stockholm in den Jahren 1943–1947.

1. Mütter, die ihre Kinder stillen, besuchten die Kinderpflegezentralen häufiger als andere Mütter. 64% der Brustkinder besuchten regelmäßig die Kinderpflegezentralen. Entsprechende Zahl für Flaschenkinder war 48%. Darum kann man kein repräsentatives Material nur durch Untersuchung der Klientel der Kinderpflegezentralen erhalten.

2. Es zeigte sich, daß die Flaschenkinder auch heutzutage die größte Sterblichkeit aufweisen. Es besteht fortwährend zwei- bis dreimal größere Todesgefahr für ein Kind, das künstlich ernährt wird, als für ein Brustkind. Die Übersterblichkeit der Flaschenkinder war am größten am Anfang des Säuglingsalters, um im vierten Quartal beinahe zu verschwinden. Weiter schienen die Kinder mit gemischter Ernährung während des zweiten und dritten Lebensmonats besonders günstig gestellt zu sein.

3. Die Übersterblichkeit der Flaschenkinder war für alle Todesursachen gültig.

4. Das Alter der Mutter und das Geburtsgewicht des Kindes schienen die Sterblichkeit bei verschiedenen Ernährungen nicht zu beeinflussen, mit Ausnahme von Prämaturen, die eine Auswahl der besonders kleinen und schwachen Brustkinder mit hoher Sterblichkeit einschlossen.

5. Der Unterschied in der Säuglingssterblichkeit zwischen verschiedenen Sozialgruppen wurde teils durch Vergleich der Kinder aus verschiedenen Einkommensstufen, teils durch Aufteilung des Materials in Hinsicht auf die Wohnungsverhältnisse beurteilt. Man

fand eine höhere Säuglingssterblichkeit in Familien mit kleinerem Einkommen und in denjenigen mit kleinerer Wohnung. Der Unterschied war jedoch früher viel größer als heute, und in den Gruppen mit höchstem Einkommen (mehr als 10,000 Schwedische Kronen jährlich) war die Säuglingssterblichkeit heutzutage dieselbe wie vor 30 Jahren. Es lag kein signifikanter Unterschied in der Verteilung der Ernährungsgruppen in den verschiedenen Einkommensgruppen vor.

6. Schließlich wurde es klar, daß weder verschiedene Einkommen noch verschiedene Ernährung die Gewichtszunahme des Säuglings merklich beeinflussen, die in der vorliegenden Untersuchung eine gute Übereinstimmung mit den von *von Sydow* früher gegebenen Ziffern zeigt.

REFERENCES

Biedert, Ph.: Die Kinderernährung im Säuglingsalter. Enke, Stuttgart 1905. – *Broman, B., G. Dahlberg* and *A. Lichtenstein:* Height and Weight during growth. Acta paediat., Stockh. *30*, 1, 1942. – *Böckh, R.:* Statistisches Jahrbuch der Stadt Berlin 1878–1917. – *Cruickshank, R.:* Infection in Infancy. Arch. Dis. Childh. *20*, 145, 1945. – *Dahlberg, G.:* Statistical Methods for Medical and Biological Students. Allen & Unwin, London 1940. – *Deeny, J.* and *E. T. Murdock:* Infant Feeding in Relation to Mortality in the City of Belfast. Brit. med. J. *1*, 146, 1944. – *Ebbs, J. H.* and *F. Mulligan:* The Incidence and Mortality of Breast and Artificially-fed. Infants admitted to Hospital with Infections. Arch. Dis. Childh. *17*, 217, 1942. – *Fröhlich, Th., A. Lichtenstein* and *S. Monrad:* Nordisk Lärobok i Pediatrik, Köpenhamn 1941. – *Grulee, C.G., H.N. Sanford* and *P.H. Herron:* Breast and Artificial Feeding. J. Amer. med. Ass. *103*, 735, 1934. – *Grulee, C.G., H.N. Sanford* and *H. Schwartz:* Breast and Artificially-fed Infants. J. Amer. med. Ass. *104*, 1986, 1935. – *Gordon, I.:* Some Social Aspects on Infant Feeding. Arch. Dis. Childh. *17*, 139, 1942. – *Gyllenswärd, C.:* Dödföddhet och tidigdödlighet i Sverige. S.O.U. 2, 1946. – *Hertz, P.:* Nogre oplysninger om de under Kopenhavns Börnetilsyn vaerende spaede Börn. Nord. hyg. tidskr. 16, 1935. – *Heubner, O.:* Reden und Abhandl. aus dem Gebiete der Kinderheilk. Barth, Leipzig 1912. – *Hjärne, U.:* Undersökningar av orsaker till sjunkande spädbarnsdödlighet i Västernorrlands och Jämtlands län 1928–1940. Upsala läkaref. förh. Ny följd. *48*, 173, 1943. – *Johannesen, A.:* Untersuchungen über den Einfluss der verschiedenen Todesursachen auf die gesamte Säuglingssterblichkeit Norwegens. Jb. Kinderheilk. *67*, 513, 1908. – *Kriege* and *Seutemann:* Ernährungsverhältnisse und Sterblichkeit der Säuglinge in Kindern. Zbl. allg. Ges. Pflege. *25*, 6, 1906. – *Lawrence, J.M., B.L. Herrington* and *L.A. Maynard:* Human Milk Studies. Comparative Values of Bovine and Human Milks in Infant Feeding. Amer. J. Dis. Childh. *70*, 193, 1945. – *Lichtenstein, A.:* Preventive Paediatries in Sweden, Acta Paediat., Stockh. *35*, 238, 1949. – *Mackintosh J.M.:* Some Observation on the Relative Imminence of Breast and Artificial Feeding and its Signifi-

cance. Proc. Nutr. Soc. *58*, 1944. – *Meyer-Delius:* cit. *Fröhlich, Lichtenstein, Monrad.* Nordisk lärobok i Pediatrik. København 1941. – *Nelson, W. E.:* Feeding of Infants. *Mitchell-Nelson:* Textbook of Pediatrics. 4th Ed. Saunders. Philadelphia 1945. – *Naumann, H.:* cit. *Fröhlich, Lichtenstein, Monrad.* Nordisk lärobok i Pediatrik. Köpenhamn 1941. – *Powers, G. F.:* Infant Feeding Historical Background and Modern Practice. J. Amer. med. Ass. *105*, 753, 1935. – *Prinzing, Fr.:* Handbuch der Mediz. Statistik, 2. Auflage. Fischer, Jena 1931. – *Rietz, E.:* Sterblichkeit und Todesursachen in den Kinderjahren. Acta paediat. Stockh. *9*, Suppl. 3, 1930. – *Rudberg, S.:* Eine Rundfrage über das Stillen in Stockholm. Acta paediat. Stockh. *24*, 92, 1939. – *Schiötz, C.:* Die Säuglingssterblichkeit in Norwegen. Acta paediat. Stockh. *15*, Suppl. 1, 1933. – *Statistisches Jahrbuch der Stadt Berlin 1878–1917.* – *Stevenson, S. S.:* The Adequacy of Artificial Feeding in Infancy. J. paediat. Stockh. *31*, 616, 1947. – *Ström, J.:* Den förebyggande mödra- och barnavården i Sverige. Nord. Med. *31*, 1560, 1946. – *Sydow, G. v.:* The Weight of Healthy Breast-fed Infants during the First Year of Life. Acta paediat. Stockh. *28*, 8, 1940. – *Thiemich, M.* and *G. Bessau:* Allgemeiner Teil, *E. Feer:* Lehrbuch der Kinderheilkunde. 10th Ed. Fischer, Jena 1930. – *Wallgren, A.:* The Neonatal Mortality in Sweden from a Paediatric Point of View. Acta paediat. Stockh. *29*, 372, 1941. – *Williams, E. J.:* Breast Feeding with Special Reference to Social Factors. Pub. Health. *60*, 201, 1947. – *Woodbury, R. M.:* Breast and Artificial Feeding Infant. Mortality and its Causes. Williams & Wilkins, Baltimore 1926.

Ettisch, G.: Acta genet. *5*, 164–179, 1955

ASCARIS SUIS ♀ IN AQUEOUS ELECTROLYTE COMBINATION

(II Contribution to "Probability Statistics in Biology")

By G. ETTISCH, New York

A. Introduction

I. The first two statistical moments of death occurrence in aqueous Rhode-Saito [20] electrolyte combination (abbreviated: RHS) will suffice to form the quantitative basis of this experimental work with Ascaris suis ♀. RHS is the medium where these beasts "normally" are kept. As results are therefore presented the corresponding "life expectancies" as well as their respective "deviation" (= [dispersion]$^{1/2}$).

In the *theoretical* part is described how the complete (probability) distribution—with unvaried exactness—was obtained under particular conditions. In the *experimental* part that method is exemplified on three different types of experiments, wanted in the further pursuit of our investigations (*Ettisch* [8]). Only the first type will have to be described in details.

II. *Ettisch* [10][1] has proved that these biologic observations fulfil all the axioms of the "collective" theory of probability evolved by *von Mises* [17], [18], [19], [19a] so that the statistics, built up on this theory may be employed in biology. It has been shown too that probability statistics are able to solve problems which the so-called "descriptive" branch was compelled to leave aside.

B. Theoretical

I. Attempting to determine the probability distribution of a "collective", a difficulty may arise:

The II. axiom of *von Mises*' theory requires

$$\mu^{\langle\nu\rangle} = \lim_{n_0 \to \infty} \mu_\nu = \lim_{n_1 \to \infty} \frac{m_\nu}{n_0} = p(x_\nu) \qquad (1)$$

[1] That paper is referred to as "Statist. I".

where m_ν is the number of appearances of the ν^{th} "attribute" x_ν in all the n_0 observations; thus $\mu_\nu = \frac{m_\nu}{n_0}$ is its "relative frequency". Equation (1) means then: If, in spite of a growing n_0, μ_ν arrives at a constant value, this empirical figure $\mu^{\langle\nu\rangle}$ is—in the defined circumstances—the probability $p(x_\nu)$ of the ν^{th} "attribute" x_ν. To obtain it n_0 actually would have to be very large[1].

But a relatively small fraction of all the observations n_0 only can fall into the initial portion of the "mixture" (v. later on) and also into the final one: "nearly all" n_0 must gravitate to the middle part of it. Thus for the shortest lifetimes as well as for the longest ones there will exist—even with a rather large n_0—only a small number of observations, maybe not sufficient to produce the limiting value of equation (1). Of course, according to axiom I (endless repeating ability of the observation) the determination of the complete probability distribution should be secured. Allowing for the asymptotic expression of axiom II as a familiar abstract manner of formulating theoretical constructions, one may nevertheless point out additionally that in any real observation such multiplying action is not only self-excluding, but is also more or less already restricted by a lot of special reasons, above all in biologic work. The principal problem remains then—and it bears clearly on our particular conditions: Is there, in a case where the complete distribution cannot be determined directly with the help of equation (1), yet a way to find each one of the ψ probabilities?

II. This general relation, employed on a defined probability distribution, is independent of any probability definition itself

$$p(x_1) + p(x_2) + \ldots\ldots + p(x_\varkappa) + \ldots\ldots + p(x_\psi) = 1 \ldots\ldots \quad (2)$$

The initial values $p(x_1)$, $p(x_2)$, ... as well as the ultimate ones $p(x_{\psi-\varrho})$, $p(x_{\psi-\varrho+\zeta})$, ... $p(x_\psi)$ of the series then will escape us. Here ϱ is any whole, positive figure of the amount: $\varrho < \psi$ so that for $\zeta = \varrho$ the result is: $p(x_\psi)$ as ζ grows from 1 to ϱ.

---

[1] As *von Mises* has put it: "The relative frequency μ_ν then has, with n_0 infinitely large, a limiting value, the probability $p(x_\nu)$ as presented in equation (1)." Some authors (*Fréchet* [12], [13]; *Kolmogoroff* [16], *Khintchine* [15]; *Cramer* [4], [5] among others) have felt bound to avoid this kind of formulation. Instead of it, it was proposed to speak of an "attribute" possessing "stochastic randomness", leading to the "true" relative frequency which finally represents the attribute's probability. *Von Mises* sees no reason for such a change. *Kamke*'s [14] proposal underlines *von Mises*' viewpoint, as do, after all, the explanations of *Copeland* [2], [3]; *Tornier* [21]; *Wald* [22], [23]; *Dörge* [6], [7], et al.

Equation (2) may be written

$$p(x_1) + p(x_2) + p(x_\varkappa) \ldots\ldots + p(x_\psi) = \sum_1^{\psi} p(x_\nu) \equiv P(x_{1,\psi}) = 1 \quad (3)$$

$P(x_{1,\psi})$ [1] is the sum of all the ψ (death) probabilities constituting the distribution.

But it may be that, for instance, only these three partial-sums, $P(x_{3,\delta})$, $P(x_{\eta,i})$, and $P(x_{\lambda,0})$ are significant out of the whole sum $P(x_{1,\psi})$, or even a possibly still greater number of partial-sums. Therefore

$$P(x_{3,\delta}) + P(x_{\eta,i}) + P(x_{\lambda,0}) < 1 \quad (4)$$

and thus

$$P(x_{3,\delta}) \equiv \sum_3^{\delta} p(x_\nu) = \lim_{n_0 \to \infty} \frac{[m_3 + m_4 + \ldots\ldots + m_\delta]}{n_0}, \quad (5)$$

then

$$P(x_{\eta,i}) \equiv \sum_\eta^{i} p(x_\nu) = \lim_{n_0 \to \infty} \frac{[m_\eta + m_{\eta+1} + \ldots\ldots + m_i]}{n_0} \quad (6)$$

and finally

$$P(x_{\lambda,0}) \equiv \sum_\lambda^{0} p(x_\nu) = \lim_{n_0 \to \infty} \frac{[m_\lambda + m_{\lambda+1} + \ldots\ldots + m_0]}{n_0} \quad (7)$$

Conditions become more simple of course, as soon as there are only two such partial sums out of the whole series of the ψ attributes, when there exists a "sum alternative". It is then spoken of the probability "*up to*" a certain point and of the probability "*greater than*" or "*beyond*" that certain point only. Equation (3) may be written

$$P(x_{1,\psi}) = P(x_{1,\gamma}) + P(x_{\gamma+1,\psi}) \equiv P(x_{1,\gamma}) + P(x'_{\gamma,\psi}) = 1 \quad (8)$$

where

$$P(x_{1,\gamma} \equiv \sum_1^{\gamma} p(x_\nu) = \lim_{n_0 \to \infty} \frac{[m_1 + m_2 + \ldots\ldots m_\gamma]}{n_0} \quad (9)$$

and

$$P(x_{\gamma+1,\psi}) \equiv P(x'_{\gamma,\psi}) \equiv \sum_{\gamma+1}^{\psi} p(x_\nu) = \lim_{n_0 \to \infty} \frac{[m_{\gamma+1} + m_{\gamma+2} + \ldots\ldots + m_\psi]}{n_0} \quad (10)$$

$P(x_{1,\gamma})$ is obviously the [sum-]probability [of death] "up to" [the time] γ, [the "mortality"] and $P(x'_{\gamma,\psi}) \equiv P(x_{\gamma+1,\psi})$ the respective

[1] The manner of designating *sum probabilities* is of necessity somewhat different from that for the *sum of single probabilities*.

[sum-]probability for a time "greater than" γ, also being called the probability of "survival" of the time γ.

From (8) follows, on the other hand:

$$P(x_{1,\gamma}) = 1 - P(x'_{\gamma,\psi}) = 1 - \sum_{\gamma+1}^{\psi} p(x_\nu) \tag{11}$$

or

$$P(x_{1,\gamma}) = 1 - \lim_{n_0 \to \infty} \frac{[m_{\gamma+1} + m_{\gamma+2} + \ldots\ldots + m_\psi]}{n_0} \tag{12}$$

where

$$\lim_{n_0 \to \infty} \frac{[m_{\gamma+1} + m_{\gamma+2} + \ldots\ldots m_\psi]}{n_0} = \lim_{n_0 \to \infty} \frac{m_{\gamma+1}}{n_0} + \lim_{n_0 \to \infty} \frac{m_{\gamma+2}}{n_0} + \ldots\ldots + \ldots\ldots + \lim_{n_0 \to \infty} \frac{m_\psi}{n_0} \tag{13}$$

Consequently, as soon as the probability "greater than" for a determined time is given, the probability "up to" the same time is calculable and vice versa without wanting further observations.

If "successive partial sum alternatives" could be formed from the beginning of the distribution to its end, making use of all of the respective "survivals", it may be possible finally, to obtain all "mortalities". This may happen in the following manner considering equations (11) and (12):

$$P(x_{1,1}) = 1 - P(x_{2,\psi}) \equiv 1 - P(x'_{1,\psi}) = 1 - \lim_{n_0 \to \infty} \frac{[m_2 + m_3 + \ldots + m_\psi]}{n_0}. \tag{14}$$

Then is

$$P(x_{1,2}) = 1 - P(x_{3,\psi}) \equiv 1 - P(x'_{2,\psi}) = 1 - \lim_{n_0 \to \infty} \frac{[m_3 + m_4 + \ldots + m_\psi]}{n_0}, \tag{15}$$

$$\vdots \qquad \vdots \qquad \vdots \qquad \vdots$$

$$P(x_{1,\varrho}) = 1 - P(x'_{\varrho,\psi}) = 1 - \lim_{n \to \infty} \frac{[m_{\varrho+1} + m_{\varrho+2} + \ldots\ldots m_\psi]}{n_0} \tag{16}$$

and finally

$$P(x_{1,\psi}) = 1 - P(x'_{\psi,\psi}) = 1 \tag{17}$$

Indeed, all the part sum probabilities of the discrete distri-

[1] This is a kind of simplified writing instead of a more complicated one. The latter is of no practical consequence and, indeed, would prevent us from recognizing clearly the nature of this particular operation.

bution seem to be thus formally obtainable by the method of "successive partial sum alternatives" described in equations (14) to (17). In the second term of the right sides of each of the equations of this system is given the respective "probability of survival".

III. But in this way a critical point Π must be reached. It belongs to (the lifetime) π attained, as soon as the number of the corresponding survivals $\sum_{\pi+1}^{\nu} m_\nu$—dwindling rapidly from this point Π to the end of the sum (curve)—is just large enough to furnish the limiting value in the manner explained in Statist. I[1]. It is

$$P(x_{1,\pi}) = 1 - P(x_{\pi,\nu}) \equiv 1 - \lim_{n_0 \to \infty} \frac{\sum_{\pi+1}^{\nu} m_\nu}{n_0} \tag{17a}$$

Reaching this critical point Π, one will be already high up on the sum curve. The help of the "survivals" may no longer be needed, for it may be possible now to calculate the mortalities $P(x_{1,\pi})$, $P(x_{1,\pi+1}) \ldots P(x_{1,\nu})$ directly from the relatively large number $\sum_{1}^{\pi} m_\nu = n_0 - \sum_{\pi+1}^{\nu} m_\nu$ etc. of the death observations at our disposal. If, however, that should not be the case, a new set of observations would have to be added to the ancient one, so that this course may be able to be continued. In order to explain this in detail, there may be for instance, the point $(\Pi + 1)$, that is, the point on the curve one "attribute" beyond that of the "critical point" Π: one arrives then at

$$P(x_{1,\pi+1}) = \lim_{n_0 \to \infty} \frac{[m_1 + m_2 + \ldots + m_{\pi+1}]}{n_0} \equiv \lim_{n_0 \to \infty} \frac{\sum_{1}^{\pi+1} m_\nu}{n_0} \tag{18}$$

$$= \sum_{1}^{\pi+1} p(x_\nu) = p(x_1) + p(x_2) + \ldots + p(x_{\pi+1}) \tag{19}$$

where obviously

$$m_1 + m_2 + \ldots + m_{\pi+1} = \sum^{\pi+1} m_\nu \tag{20}$$

is the total number of death observations up to the actual lifetime $(\pi + 1)$ of point $(\Pi + 1)$.

[1] This number was then determined for one single, particular value of the relative frequency only, for the average lifetime, in order to exemplify the procedure. As soon as the whole distribution has to be determined a considerably larger number of observations will be needed.

The corresponding number of "survivors" is

$$\sum_{\pi+2}^{\psi} m_\nu = m_{\pi+2} + m_{\pi+3} + \cdots\cdots + m_\psi, \tag{21}$$

but their relative frequency is

$$\mu_{(\pi+2)} = \frac{\sum_{\pi+2}^{\psi} m_\nu}{n_0} \neq \lim_{n_0 \to \infty} \left| \frac{[m_{\pi+2} + m_{\pi+3} + \cdots\cdots + m_\psi}{n_0} \right| \neq P(x_{\pi+2,\psi}) \tag{22}$$

meaning: here does not exist the limiting value of the relative frequency the probability that is, because of lack of observations.

IV. Evidently, one finds now from the lifetimes v and $(v-1)$:

$$P(x_{1,v}) - P(x_{1,v-1}) = p(x_v) \tag{23}$$

where

$$v = 1, 2, \ldots\ldots \psi \tag{24}$$

Equation (23) is our essential result. Any single probability $p(x_\varkappa)$ may now be understood as

$$p(x_\varkappa) = P(x_{1,\varkappa}) - P(_{1,\varkappa-1}) = \sum_{1}^{\varkappa} p(x_\nu) - \sum_{1}^{\varkappa-1} p(x_\nu) \tag{25}$$

That is–considering (1)—indeed

$$p(x_\varkappa) = \lim_{n_0 \to \infty} \frac{m_\varkappa}{n_0} \tag{26}$$

Two margin values follow at once from (25). The first is the single probability for $\varkappa = 1$, the initial attribute,

$$p(x_1) = P(x_{1,1}) \tag{27}$$

Using equation (14) it may be written, according to (8) and (10):

$$p(x_1) = 1 - P(x'_{1,\psi}) \equiv 1 - \left[p(x_2) + p(x_3) + \cdots\cdots + p(x_\psi)\right], \tag{28}$$

or

$$= 1 - \sum_{2}^{\psi} p(x_\nu); \tag{29}$$

and correspondingly the last single probability $p(x_\psi)$

$$p(x_\psi) = 1 - P(x_{1,\psi-1}) = 1 - \lim_{n_0 \to \infty} \frac{[m_1 + m_2 + \cdots\cdots + m_{\psi-1}]}{n_0}. \tag{30}$$

It may now be deduced

$$p(x_2) = P(x_{1,2}) - p(x_1) = \sum_{1}^{2} p(x_\nu) - p(x_1) \tag{31}$$

. . .
. . .
. . .
. . .
. . .

$$p(x_\varepsilon) = P(x_{1,\varepsilon}) - P(x_{1,\varepsilon-1}) = \sum_{1}^{\varepsilon} p(x_\nu) - \sum_{1}^{\varepsilon-1} p(x_\nu) \tag{32}$$

.
.
.
.

and finally

.
.

$$p(x_\psi) = 1 - P(x_{1,\psi-1}) = \sum_{1}^{\psi} p(x_\nu) - \sum_{1}^{\psi-1} p(x_\nu) \tag{33}$$

identical with (30).

Thus the complete arithmetic distribution of our ψ probabilities is obtained.

For the rest it is clear from (2) or from (4) that, in order to know the entire respective distribution we need know empirically only any $(\psi-1)$ of the ψ constituting probabilities. The missing $p(x_\chi)$ may then be calculated

$$p(x_\chi) = 1 - \left| \sum_{1}^{\chi-1} p(x_\nu) + \sum_{\chi+1}^{\psi} p(x_\nu) \right| \tag{34}$$

V. There are, of course, other methods to find the discontinuous distribution, for instance such ones securing the distribution function. Here a relatively small number of observations is given and the (geometric) distribution is calculated. Such are the interpolation methods of *Newton*, *Lagrange*, and others, as well as a particular *von Mises* method [17], essentially like an older one of *Bruns* [1]. The *von Mises* method has the great advantage of furnishing an exponential function closely related to the *Laplace-Gauss* and to the *Maxwell-Boltzmann* familiar distribution functions. The method proposed here seems particularly convenient; for with the other ones the discrete (arithmetic) distribution has to be transformed first into a continuous (geometric) one which may bring with it lengthy calculations. Moreover, our procedure furnishes a series of other useful figures besides the $p(x_\nu)$. It is this method of the "successive partial sum alternatives" that has been employed in the following investigations.

C. Experimental

I. The experiments, carried out with Ascaris suis ♀, were of three types:

Type 1 (abbreviated: T_1). The first two statistical moments of the lifetime of whole, free Ascaris in RhS. were determined.

Type 2 (T_2). The cephalic segment, about 4.0 cm long, cut from the trunk, was observed in the manner already described in previous papers (*Ettisch* [8], [10]).

Type 3 (T_3). The cephalic segment again was observed; this time completely free in all its movements.

The Ascaris came from the slaughterhouse to the laboratory in thermosbottles filled with RhS. at 310.8/311.8° K. immediately after the hog was killed. The electrolyte was at once replaced by a fresh one of the same temperature. Only animals of apparently equal shape, mobility, coloration, etc. were chosen out of a great number. Further principal details of the procedure have already been given (*Ettisch* [l.c.]).

II. The observations ("collective") of T_1 are found in Table I.

Table I. Original Collective (corrected).

Type I (hours)						
335	165	287	165	386	76	406
261	357	386	165	165	549	335
312	452	237	192	165	431	142
192	549	165	431	357	528	142
142	549	261	406	192	312	192
76	357	192	357	431	165	287
215	605	215	215	406	261	312
386	192	119	215	476	165	335
142	192	237	192	335	119	364
119	142	335	237	357	215	406
165	287	431	386	406	237	406
192	237	386	76	287	165	573
165	237	312	142	192	142	431
237	215	43	192	215	215	142
237	237	43	261	192	215	215
215	335	192	192	215	165	237
165	312	287	287	119	119	237
192	312	287	335	142	287	261
215	335	335	237	237	192	165
192	335	119	287	165	261	237
165	312	549	528	431	287	

In the first, raw notations the "time of observation" began with the hour the Ascaris were selected in the laboratory. It ended at the hour the animal was found dead. Between these two events checks were made early every morning and every evening, often a third time about midnight.

This involves a certain error. The animals could have died before the time of the respective observations, or they may only have seemed dead on inspection but may not actually have been so. One appeared justified in assuming possible differences of $\pm$ 6 hours, constituting an error of about 15 % for the smallest lifetime and of 1.2 % for the highest. As for the mean values, the error is then 2.3 % and 3.9 % for the highest and lowest figure, respectively. It will be possible to increase the accuracy of the observations by introducing finer methods[1].

In the T_2 experiments the error should be somewhat lower; nevertheless, the same figure was chosen for these observations. Thus all raw notations for T_1, T_2 and T_3 were adjusted to an error of $\pm$ 6 hours.

There is another point still to be mentioned.

The "collective" theory of probability starts from a set of figures that demonstrates completely the character of contingency (III axiom)[2]. An example of this contingency is represented by the velocities of the molecules of an ideal gas, as they are found under determined conditions of state. That gives just the "original" collective C_o, after having considered the corresponding error.

If, however, an order was introduced, for instance if the random velocities were arranged in an ascending scale, then, according to *von Mises*, one of the "elementary statistical operations" was carried out, the so-called "mixture", obtaining the collective C_m out of C_o. The lifetime of our Ascaris was *not* recorded in the order the beasts died. Having selected, from a great number, animals as nearly alike as possible, to each of them in the same fortuitous manner was assigned its respective individual lifetime τ_ν. Thus the original collectives $C_o^{T_1}$, $C_o^{T_2}$ and $C_o^{T_3}$ respectively were got.

These three sets of figures actually fulfil all three axioms of the "collective" theory[3].

Thus we arrived at a result that would have been obtained also had we been compelled—as one is nearly always in physics—to make only one observation at a time.

Having carried out with the original collectives $C_o^{T_1}$ etc. the so called "mixture" the ensueing collectives $C_m^{T_1}$, $C_m^{T_2}$, $C_m^{T_3}$, show necessarily that ascendant sequence of figures which is often believed to represent the only possible kind of observation[4].

III. Table II column 1 gives the number of the ν probabilities ("attributes") composing the distribution. In column 2 are the ν

[1] We succeeded in essentially improving the exactness of the later measurements by introducing a series of helpful contrivances.

[2] We assumed axiom III fulfilled when the conditions *von Mises* prescribed were fulfilled.

[3] In Statist. I it was circumstantially explained how this can be shown.

[4] Experience of a strange kind has taught us the necessity for the discussion of the trivial matter contained in the preceding five paragraphs.

death attributes τ_ν themselves in the accustomed ascendant order. In column 3 are found the "mortalities" from $P(\tau_{1,1})$ to $P(\tau_{1,23})$; column 4 gives the corresponding ψ single probabilities, the $p(\tau_\nu)$.

Table II. Distribution

Type I			
No	τ_ν (attributes)	$P(\tau_{1,\nu})$ (mortalities)	$p(\tau_\nu)$ (single probabilities)
	hours		
1	$\tau_1 = 43$	$P(\tau_{1,1}) = 0.01$	$p(\tau_1) = 0.01$
2	$\tau_2 = 76$	$P(\tau_{1,2}) = 0.03$	$p(\tau_2) = 0.02$
3	$\tau_3 = 119$	$P(\tau_{1,3}) = 0.08$	$p(\tau_3) = 0.05$
4	$\tau_4 = 142$	$P(\tau_{1,4}) = 0.14$	$p(\tau_4) = 0.06$
5	$\tau_5 = 165$	$P(\tau_{1,5}) = 0.25$	$p(\tau_5) = 0.11$
6	$\tau_6 = 192$	$P(\tau_{1,6}) = 0.36$	$p(\tau_6) = 0.11$
7	$\tau_7 = 215$	$P(\tau_{1,7}) = 0.45$	$p(\tau_7) = 0.09$
8	$\tau_8 = 237$	$P(\tau_{1,8}) = 0.55$	$p(\tau_8) = 0.10$
9	$\tau_9 = 261$	$P(\tau_{1,9}) = 0.59$	$p(\tau_9) = 0.04$
10	$\tau_{10} = 287$	$P(\tau_{1,10}) = 0.66$	$p(\tau_{10}) = 0.07$
11	$\tau_{11} = 312$	$P(\tau_{1,11}) = 0.71$	$p(\tau_{11}) = 0.05$
12	$\tau_{12} = 335$	$P(\tau_{1,12}) = 0.77$	$p(\tau_{12}) = 0.06$
13	$\tau_{13} = 357$	$P(\tau_{1,13}) = 0.81$	$p(\tau_{13}) = 0.04$
14	$\tau_{14} = 364$	$P(\tau_{1,14}) = 0.82$	$p(\tau_{14}) = 0.01$
15	$\tau_{15} = 386$	$P(\tau_{1,15}) = 0.85$	$p(\tau_{15}) = 0.03$
16	$\tau_{16} = 406$	$P(\tau_{1,16}) = 0.89$	$p(\tau_{16}) = 0.04$
17	$\tau_{17} = 431$	$P(\tau_{1,17}) = 0.93$	$p(\tau_{17}) = 0.04$
18	$\tau_{18} = 452$	$P(\tau_{1,18}) = 0.94$	$p(\tau_{18}) = 0.01$
19	$\tau_{19} = 476$	$P(\tau_{1,19}) = 0.95$	$p(\tau_{19}) = 0.01$
20	$\tau_{20} = 528$	$P(\tau_{1,20}) = 0.96$	$p(\tau_{20}) = 0.01$
21	$\tau_{21} = 549$	$P(\tau_{1,21}) = 0.98$	$p(\tau_{21}) = 0.02$
22	$\tau_{22} = 573$	$P(\tau_{1,22}) = 0.99$	$p(\tau_{22}) = 0.01$
23	$\tau_{23} = 605$	$P(\tau_{1,23}) = 1.00$	$p(\tau_{23}) = 0.01$

In every single $P(\tau_{1,\nu})$ or $P(\tau'_{\nu,\psi})$ there was checked whether the respective number of survivors—or eventually deaths—really furnished the limiting value required by the theoretical explanations of Statist. I. The same was done for T_2 and T_3.

IV. Table III, column 3 gives the life expectancy of the three types with their—fairly small—mean error.

The freely moving cephalic segments (T_3) thus have the same lifetime as the whole, free living, uninjured animals (T_1), within the limits of experimental error. The segments in the recording device (T_2), however, have a more than 40% smaller life expectancy.

Columns 4 and 5 give the mortalities "up to" the respective life expectancies $P(\tau_{1,\bar{\tau}})$ and for each of the corresponding probability "greater than" the expectancy $P(\tau_{\bar{\tau},\psi})$, the "probability of survival".

The average of the 3 figures of column 4 is $P(\tau_{1,\bar{\tau}}) = 0.58 \pm 0.024$ ($\pm 4.1\%$) and of those of column 5 is $P(\tau_{\bar{\tau},\psi}) = 0.42 \pm 0.024$ ($\pm 5.7\%$). For all three types of experiments may thus be assumed the same mortality up to the life expectancy and, consequently, a common figure for the corresponding survivals.

If a sufficiently large number of Ascaris in RhS is observed, it is of no importance in what state the animals are encountered, whether in T_1, T_2, or T_3. At average lifetime the number of the survivors constitute 42%, while the death rate is 58%. At the mean lifetime 16% more beasts will thus have died than will survive.

V. Column 6 gives the number n_o^* of observations necessary to reach the limiting value of the relative frequency of the respective average lifetime for the three types of experiments (cf. Statist. I).

Table III

No	Type of Experiment	Expectancy value	Mortality	Survival		Deviation
		$\bar{\tau} \times 10^{-2}$	$P(\tau_{1,\bar{\tau}})$	$P(\tau'_{\bar{\tau},\psi})$	n_o^*	$\bar{s}$
1	2	3	4	5	6	7
		hours				*hours*
1	T_1	$2.65 \pm 4\%$	0.59	0.40	60 Exp.	$116 \pm 10\%$
2	T_2	$1.53 \pm 6\%$	0.54	0.46	30 Exp.	$72 \pm 10\%$
3	T_3	$2.47 \pm 4\%$	0.62	0.38	40 Exp.	$76 \pm 10\%$

It seems significant that one cannot rely on a generally fixed number of observations in order to arrive at a result, independent of the number of observations. To assume otherwise, proposed by authors working with descriptive statistics, is to ignore the fact that this important figure will depend on the internal state of the beasts and this again will vary, as was shown by *Ettisch* (l.c.), with external conditions.

VI. Finally, column 7 gives the deviation

$$\bar{s} = \left| \sum_{1}^{\psi} (\bar{\tau} - \tau_\psi)^2 \, p(\tau_\psi) \right|^{\frac{1}{2}} \tag{35}$$

$\bar{s}_{T_1}$ shows the highest value, while $\bar{s}_{T_2}$ and $\bar{s}_{T_3}$ are equal within the limits of experimental error. It may be remembered that T_1 is

considered as the "normal" state of the animal—an untouched Ascaris suis ♀ in the "keeping" medium.

VII. The life expectancies for T_1 and T_3 (τ_{T_1} and τ_{T_3}) are equal, but the deviation $\bar{s}_{T_3}$ is only about 65% of s_{T_1}. That means, only the average inequalities of the respective n_0 life times for T_1 and T_3 differ. In other words: The area on which the n_0 single life times of T_1 and T_3 are scattered with respect to τ_{T_1} (or τ_{T_3})—their gravitational center—is retracted by the corresponding amount. Therefore, cutting off of the cephalic segment displaces the single life times of T_3—opposite those of T_1—so that $\bar{\tau}_{T_3}$ remains equal to τ_{T_1}, but $\bar{s}_{T_3}$ is only $0{\cdot}65\,\bar{s}_{T_1}$. The sum curve of T_3 is shrunk, compared with that of T_1, and has rotated, by a perpendicular axe through the gravitational center τ_{T_1} in the direction opposed to that of the clock-hands.

The average duration of the life determining processes (metabolism, circulation, buffer reactions et al.) has not changed after the cutting off of the cephalic segment, but the statistical measure of the difference of their single durations has become smaller—the mass center $\bar{\tau}_{T_3}$ being the reference point.

If, however, the cephalic segment is put to work in the recording device, (T_2), τ_{T_2} as well as s_{T_2} turn smaller by more than 35%. The duration of the life determining processes of the organism are then, in the average, shortened by that amount and so is the difference in the durations, compared with those of whole, uninjured animals.

$\bar{s}_{T_2}$ is statistically the same as that for T_3.

VIII. (1) Table I shows very exactly the number total of observations n_0 necessary and sufficient to obtain the whole distribution.

In the first place n_0^* for the τ_{T_1} point was determined according to Statist. I. Then the whole was treated as described above. Thus is substituted by a rational method the claim for larger and larger numbers of experiments, where scientific reasoning indicates neither the need for, nor any advantage from it.

(2) Table II col. 3 shows at its end

$$P(\tau_{1;23}) = 1 \tag{36}$$

This proves that the observations of T_1 (as also those of T_2 and T_3) fulfil the fundamental theorem of the starting point, namely of

equation (2) or (3), as soon as they are treated in the way proposed in BI, BII and BIII of this paper.

(3) From equation (36) follows also that the discrete links, constituting that sum must be probabilities. That means: The second term of the right side in each of the equations (14) to (17) is not any possible relative frequency with an amount depending on n_o, but is really the constant limiting value of it, independent of n_o,—the probability.

(4) The relation

$$\sum_{1}^{23} p(\tau_\nu) = 1 \tag{37}$$

represents, according to equation (3), a statement identical to that of (36). Nevertheless, it has its special importance insofar as—taking the figures of Table II col. 4—the sum of just these ν discrete values of the probability distribution produces the expected unit. This constitutes a special confirmation for the proposed method.

Equation (37) thus may give added confidence in the calculation of the single probabilities of the distribution according to the proposed method and at the same time it may serve as a proof against error in the calculating itself.

(5) Table III cols. 4 and 5 give for a special pair of values of our "sum alternatives", for those concerning the life expectancy $\bar{\tau}$

$$P(\tau_{1,\bar{\tau}}) + P(\tau_{\bar{\tau},\nu}) = 1 \tag{38}$$

for all 3 types of experiments. Both terms of that sum (38) had been calculated independently from each other. The situation of $\bar{\tau}_{T_1}$, $\bar{\tau}_{T_2}$, $\bar{\tau}_{T_3}$ permitted that.

I like to remember the correspondence concerning this paper I once had with Dr. R. Fuerth, Edinburgh.

My thanks go to S^ra *D. Elisa Marques de Silva*, Instituto Portugues de Oncologia, Lisbon-Palhava, for her indefatigable help with the tiresome numerical calculations of this paper and the previous one.

Summary

I. Using *von Mises*' probability statistics in three different kinds of experiments with Ascaris suis ♀ in electrolyte combination the problem arises: How is it possible to find the probabilities not only for the middle part of the—discontinuous—distributions, but also for their beginnings and their ends, without being compelled to resort to too high a number of observations?

II. The solution here proposed consists:
First, in determining all the partial-sum probabilities, the "mortalities", from $P(\tau_{1,1})$ to $P(\tau_{1,\psi})$. This becomes possible by forming "successive partial-sum alternatives", using in this way also the "probabilities of survival".

The conditions are described under which this method of the "successive partial-sum alternatives" will work.

Second, in determining the (discontinuous) distribution of the ψ $p(\tau_\nu)$ probabilities by a simple arithmetic operation with the "mortalities" $P(\tau_{1,\psi})$. The admissibility of the procedure has been proved.

III. This method has been employed in our three types of observations. Expectancy values as well as deviations are calculated and the conclusions drawn.

Résumé

I. S'étant servi des statistiques de probabilité de *von Mises* dans trois types différents d'expériences poursuivies avec l'ascaris suis femelle élevée dans une solution combinée d'électrolytes, l'auteur soulève le problème suivant: comment est-il possible de déterminer les probabilités, non seulement de la partie moyenne des distributions (discontinues) mais aussi de leur début et de leur fin, sans être contraint d'utiliser un nombre par trop élevé d'observations?

II. La solution suivante est proposé: Premièrement, en déterminant toutes les probabilités sommatoires partielles, les «mortalités», de $P(\tau_{1,1})$ à $P(\tau_{1,\psi})$. Cela devient possible grâce à la formation d'«alternatives sommatoires partielles successives», se servant ainsi également des «probabilités de survivance».

On décrit les conditions dans lesquelles cette méthode d'«alternatives sommatoires partielles successives» peut être appliquée. Deuxièmement, en déterminant la distribution discontinue des caractères $\psi p(\tau_\nu)$ par une opération arithmétique simple, basée sur les «mortalités» $P(\tau_{1,\psi})$. L'auteur démontre que ce procédé est admissible.

III. Cette méthode fut alors employée pour l'étude des trois types d'observations considérés. L'auteur calcula les valeurs probables ainsi que les déviations avant de tirer les conclusions de son travail.

Zusammenfassung

I. Bei der Verwendung der Wahrscheinlichkeitsstatistik von *von Mises* für drei verschiedene Arten von Versuchen mit Ascaris suis ♀ in Elektrolysenverbindung taucht folgendes Problem auf: Wie kann man die Wahrscheinlichkeiten nicht nur für den mittleren Teil der – diskontinuierlichen – Verteilungen, sondern auch für ihre Anfänge und Endteile finden, ohne gezwungen zu sein, auf eine zu große Anzahl von Beobachtungen zurückgreifen zu müssen?

II. Die hier vorgeschlagene Lösung besteht: 1. In der Bestimmung aller partiellen Summenwahrscheinlichkeiten, der «Sterblichkeiten», von $P(\tau_{1,1})$ bis $P(_{1,\psi})$. Dies wird möglich durch die Bildung von «sukzessiven Teilsummen-Alternativen», wobei man auf diesem Wege auch die «Wahrscheinlichkeiten des Überlebens» verwendet. Die Bedingungen, unter denen diese Methode der «sukzessiven Teilsummen-Alternativen» funktioniert, sind dargelegt. 2. In der Bestimmung der diskontinuierlichen Verteilung der $\psi p(\tau_\nu)$ Eigenschaften durch eine einfache, arithmetische Operation mit den «Sterblichkeiten» $P(\tau_{1,\psi})$. Die Zulässigkeit des Verfahrens wird bewiesen.

III. Diese Methode wurde dann bei unseren drei Typen von Beobachtungen angewendet. Erwartete Werte sowohl als auch Abweichungen wurden berechnet und die Schlußfolgerungen gezogen.

REFERENCES

1. *Bruns, H.:* Wahrscheinlichkeitsrechnung und Kollektivmaßlehre. Leipzig, Berlin 1906. – 2. *Copeland, A.H.:* Amer. J. Math. *58*, 181, 1936.– 3. *Id.:* Lectures No. 4, Notre Dame Univ. – 4. *Cramér, H.:* Random Variables and Probability Distributions. Cambridge University Press 1937. (Cambridge Tracts in Mathematics and Mathematical Physics No. 36). – 5. *Id.:* Mathematical Methods of Statistics. Princeton 1946. – 6. *Dörge, K.:* Jb. Dtsch. math. Vereinig. *43*, 39, 1934. – 7. *Id.:* Math. Z. *32*, 232, 1930; *40*, 161, 1935. – 8. *Ettisch, G.:* J. gen. Physiol. *34*, 43, 1950 (with detailed bibliography). – 9. *Id.:* will appear shortly. – 10. *Id.:* Arq. Pat. *17*, 2, 1945. – 11. *Id.:* Fehlerrechnung in *P. Ronas*' Repetitorium der Physiologischen Chemie, 2. Bd., p. 694. J. Springer, Berlin 1929. – 12. *Fréchet, M.:* The Diverse Definitions of Probability. In *Borel, E.F.E.J.:* Traité du Calcul des Probabilités, vol. 1, fasc. 3. Gautier-Villars, Paris. – 13. *Id.:* Exposé et Discussion de Quelques Recherches Récentes etc. Théories des Probabilités. Fasc. II Actualités Scientifiques et Industrielles. No 735, Paris 1938. – 14. *Kamke, E.:* Einführung in die Wahrscheinlichkeitstheorie. S. Hirzel, Leipzig 1932, 2 ff. – 15. *Khintchine, A.:* Asymptotische Gesetze der Wahrscheinlichkeitsrechnung. Ergebnisse der Mathe-

matik und ihrer Grenzgebiete. J. Springer, Berlin 1933. – 16. *Kolmogoroff, A.:* Grundbegriffe der Wahrscheinlichkeitsrechnung. J. Springer, Berlin 1955. – 17. *Mises, R. von:* Wahrscheinlichkeitsrechnung. F. Deuticke, Leipzig/Wien 1931. – 18. *Id.:* Wahrscheinlichkeit, Statistik und Wahrheit. J. Springer. Wien 1936. – 19. *Id.:* English edition of foregoing book. Macmillan, London 1939. – 19a. *Id.:* Ann. Inst. *H. Poincaré 3*, 137, 1932. – 20. *Rhode, H.* and *Yutaka Saito:* Z. ges. exp. Med. *38*, 503, 1923. – 21. *Tornier, E.:* Wahrscheinlichkeitsrechnung und allgemeine Integrationstheorie 1936. – 22. *Wald, A.:* Sequential Analysis. p. 8. John Wiley and Sons, Inc., New York 1947. – 23. *Id.:* 4 Lectures. 1941, p. 2. Notre Dame, Ind. 1942.

LIBRI

H. Schade: **Vaterschaftsbegutachtung (Investigations of questioned paternity).** Schweizerbartsche Verlagsbuchhandlung, Stuttgart 1954. 250 Seiten.

Every physician and geneticist is by now well acquainted with the use of the different inherited serological traits in legal medicine *e.g.* in cases of questioned paternity. This method permits an exclusion of approximately 45 per cent of *falsely* accused fathers. It is regulated by law and currently used in a large number of countries.

While the serological traits have been recognized only as a method of exclusion additional methods have been developed to furnish positive evidence of paternity. These procedures have been based on available experiences in human and medical genetics and have been called anthropologic-genetic examinations. Originally developed in Austria and Germany, where it has been legal some 20 years, the anthropologic-genetic method has recently also been recognized and regulated by law in a few other European countries *e.g.* in Sweden in 1950.

By and large the anthropologic-genetic examination is a method of polysymptomatic similarity diagnosis similar to that used in the diagnosis of monozygotic twins. Although this method generally is incapable of furnishing definite proof or disproof of paternity one now feels convinced that it has a very great value for the courts.

Dr *Schade's* book is the first comprehensive treatise on this subject and should therefore be welcomed by all human geneticists concerned with paternity law suits. Also the physical anthropologist and the sociologist will find much of interest here. The evaluation of different procedures and currently used genetic traits is critical and careful. The anthropologic-genetic method does by no means represent a finished and fool-proof procedure and those who hope to find some simple rules of thumb will be dissappointed. Much research work remains to be made in order to increase the reliability of the method. Nevertheless in the hands of experienced human geneticists the method is now of great practical importance.

Jan A. Böök, Uppsala